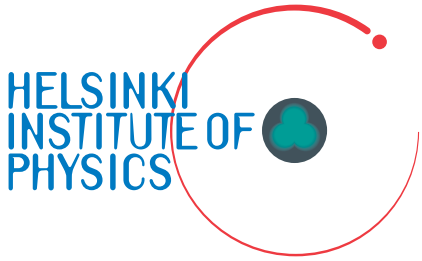


HIP Internal Report Series
HIP-2017-03

Precision measurements of $\mathrm{B}_{\mathrm{s}}^{0}$ mesons in the CMS experiment at the LHC

Terhi Järvinen

Department of Physics
Faculty of Science
UNIVERSITY OF HELSINKI

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Science of the University of Helsinki, for public criticism in the auditorium E204 at 10 o'clock, on December 18th, 2017.

Helsinki 2017

Supervisors

Professor Paula Eerola, University of Helsinki

Dr. Giacomo Fedi, University of Pisa and INFN Pisa

Pre-examiners

Reader Lars Eklund, University of Glasgow

Dr. Weiming Yao, Lawrence Berkeley National Laboratory

Opponent

Professor Ulrik Egede, Imperial College London

ISSN 1455-0563

ISBN 978-951-51-1271-2 (paperback)

Helsinki University Print (Unigrafia Oy)

ISBN 978-951-51-1272-9 (pdf)

`http://ethesis.helsinki.fi`

Electronic Publications at the University of Helsinki

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Laitos — Institution — Department
Faculty of Science	Department of Physics

Tekijä — Författare — Author
Terhi Järvinen

Työn nimi — Arbetets titel — Title

Precision measurements of B_s^0 mesons in the CMS experiment at the LHC

Oppiaine — Läroämne — Subject
Physics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
PhD Thesis	November 2017	181 pages

Tiivistelmä — Referat — Abstract

Two precision measurements are performed to B_s^0 data. The primary measurement of the thesis is the effective lifetime of the B_s^0 meson decaying in the $J/\psi\phi(1020)$ state. The effective lifetime analysis uses 2012 data recorded by the Compact Muon Solenoid (CMS) experiment. The data are collected with a center-of-mass energy of 8 TeV and correspond to an integrated luminosity of 19.7 fb^{-1}. The effective lifetime is found to be $c\tau_{eff} = 443.9 \pm 2.0$ (stat.) ± 1.2 (syst.) μm. In addition to the lifetime measurement, the design, validation and performance of the two flavour tagging algorithms are described. The algorithms are developed to improve the precision of the weak mixing phase measurement. Lastly, the results from the weak phase analysis are briefly covered. The flavour tagging studies and the weak phase measurement are also performed with the 2012 data recorded by the CMS experiment.

Avainsanat — Nyckelord — Keywords
B physics, b-hadron lifetime measurements, CP violation

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

Contents

Preface

In this thesis, neutral B_s^0 meson system is studied using the data collected by the CMS experiment. In the first part of my PhD degree, I designed and optimised the first flavour tagging algorithm used in the early weak phase measurement of the $B_s^0 \to J/\psi\phi$ decay mode. I also contributed to validation of the second, more advanced tagging algorithm that was created at INFN and the University of Padua and used in the CMS run-I paper of the weak mixing phase analysis. The latter part of my degree was devoted to an effective lifetime measurement using the $B_s^0 \to J/\psi\phi$ data. The lifetime measurement was done in the context of the CMS run-I legacy paper of the b-hadron lifetimes. From these two analyses, my contribution to the effective lifetime measurement is more significant, since I was the leading analyst of the measurement.

Along with the data-analysis tasks, I have done experimental work at the CMS experiment by monitoring the quality of the collected data at the CMS control room. I was also involved in the tracker alignment studies. I participated in the data certification process of the level-1 trigger subsystem, whose functioning is crucial for successful data collection. In parallel with the detector performance duties, I collaborated with the level-1 trigger group by coordinating, updating and maintaining the instructions related to the performance monitoring of the level-1 trigger system.

Acknowledgements

The research documented in this thesis was done at Helsinki Institute of Physics (HIP) and at the Compact Muon Solenoid (CMS) experiment at the European Organization for Nuclear Research (CERN). The research was funded jointly by the HIP CMS project, and the Magnus Ehrnrooth foundation. Travels to schools, conferences, and CERN have been funded by HIP, the doctoral programme of Particle and Universe Sciences at the University of Helsinki, as well as the Oskar Öflund foundation.

I would like to thank professor Paula Eerola for giving me an opportunity to work in her group and introducing me an inspiring world of B physics. I wish to thank my supervisor and collaborator Dr. Giacomo Fedi for his support, encouragement and highly useful ideas regarding many of the most difficult challenges in the analyses I have been involved in. I have learnt a lot from you and I am very happy that we worked together. During the effective lifetime measurement, I also worked with two excellent young scientists, Dr. Sara Fiorendi and Dr. Jhovanny Andres Mejia Guisao. Thank you for fruitful collaboration and all the hard work done for the sake of the b-hadron lifetime paper.

A big thanks goes to emeritus professors Jorma Tuominiemi and Paul Hoyer for their comments, ideas and improvements related to the thesis manuscript. Without your help, the thesis would not have the same quality as it now has. I am also very grateful for my thesis examiners Dr. Weiming Yao and Dr. Lars Eklund who

steadily read through the thesis and gave their statements and improvements in a short amount of time.

I would like to thank present and previous conveners of the CMS B physics group, Dr. Martino Margoni and Dr. Fabrizio Palla, for their guidance and help during the research projects I have been working on. I would also like to thank Dr. Tristan du Pree, whose help was crucial in completing the apparent efficiency studies related to the effective lifetime measurement.

A special thanks goes to the superb HIP-CMS gang of doctoral researchers Jaana, Juska, Joona, Santeri and Hannu. It has been very nice to hang out, organise PR events together and share lunch and coffee breaks with you. I would also like to thank the HIP staff, Tapio, Mikko, Tomas, Sami, Matti and Kati for the help given in practical, technical and administrative issues related to the PhD degree, research and PR activities.

During my one year visit at Geneva, I was privileged to get to know a group of great CERN folks with whom I shared an apartment and who became my family abroad. Thank you Arabella, Jui-Fa, Craig and Hideyuki for the cultural exchange and the rest of the jazz organized in our flat. At CERN, I also got to know Morgane and Lukasz, whose cheerfulness lightened up my last months in Geneva. A special mention goes to Maria, the renowned founder of the CERN Uterus Network Team which activities I also got to participate in.

In countless activities of Resonanssi ry, the student union of physics students at the University of Helsinki, I have been lucky to befriend with absolutely distinctive people and experience student years with them. They have also made me feel the department of physics my intellectual home. A heartfelt thanks goes to the mind-blowing freshman girls from the academic year 2007: Sanna, Sanni, Sofia, Maija, Harriet, Iida, and Heidi. It has been great to grow up and share these years with you. My student years at Kumpula have been filled with a great deal of fun and hard

work in studies and at Resonanssi, that I have experienced with Olli, Olli-Pekka, Mikko, Pekka, Reko, Petri, Markku, Laura, Daniel, and Jari. Thank you for the cheerful moments passed and yet to come.

Finally, I wish to thank my parents Kirsi and Markku, as well as, my sister Marjut for their encouragement and love. Last but not least, my deepest gratitude goes to my dear partner Alessandro, for his love, support, and understanding.

Helsinki, November 2017

Terhi Järvinen

1. Introduction

The beauty quark is the heaviest quark able to form hadrons, and hence it plays a special role in heavy-flavour physics. Beauty hadrons are characterised by a b or a $\bar{\text{b}}$ quark accompanied by much lighter up, down, strange or a charm quark. Large numbers of b hadrons are produced at the Large Hadron Collider and their distinct decay signatures can be recorded and identified by the Compact Muon Solenoid detector [1].

A large variety of high-precision measurements offer numerous opportunities to study weak decays of b hadrons. One notable branch of precision analyses are lifetime measurements. The lifetime measurements of b hadrons are interesting since they help to shed light on the bound-state dynamics of these particles. Precise measurements of b-hadron lifetimes are important in probing the standard model (SM) of particle physics - the most accurate theory to date describing the observed behaviour of the constituents of matter. A hot topic in flavour physics is the observed charge-conjugation parity (CP) violation within the quark sector. CP violation in the SM is orders of magnitude too small to explain the dominance of matter in the universe. This suggests that there must be undiscovered sources of CP asymmetry that are likely related to undiscovered new physics.

Neutral B_s^0 ($\bar{\text{b}}\text{s}$) and B_d^0 ($\bar{\text{b}}\text{d}$) mesons have many properties that make them an ideal subject for lifetime and CP violation studies. The B_s^0 and B_d^0 mesons are known to have two eigenstates with different decay widths and masses. The decay

width difference of the two eigenstates is particularly large in the B_s^0 meson system while for the B_d^0 meson the width difference is close to zero [2]. Both particles can also oscillate into their antiparticles and decay into final states that are accessible for both B and $\bar{B}$ mesons. Examples of such decay modes are $B_s^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{0^*}$.

A theory framework describing many aspects of b hadron decays is called Heavy Quark Effective Theory and Heavy Quark Expansion (HQE) [2]. In the HQE, the decay rates of b hadrons can be expressed as a sum of a series of the operator expectation values having increasing dimension. To the first order of the expansion, the b hadron decay is entirely driven by the flavour-changing weak transition $\bar{b} \to W\bar{q}'$ while the light spectator quark (u,d,c,s) inside the hadron stays intact. For this reason, the b hadrons are predicted to have the same lifetime to the first order of the expansion [2]. In reality, several effects change the lifetimes of b hadrons up to 10%. Such mechanisms include weak scattering, annihilation, and Pauli interaction as well as interference of decay amplitudes [3]. Lifetime measurements of the different b hadrons help to determine the importance of these effects in the b sector.

CP violation can be studied in the $B_s^0 \to J/\psi\phi$ channel by measuring the weak mixing phase ϕ_s of the B_s^0 - $\bar{B}_s^0$ system. The weak phase is a parameter related to the angle β_s of a unitarity triangle determined from the CKM matrix. The standard model predicts a precise and small value for the weak phase $\phi_s \approx -2\beta_s = -0.0370 \pm 0.0006$ rad [4]. If the measured ϕ_s deviates from the SM prediction, this could indicate the presence of new physics.

The weak phase and lifetimes of the B_s^0 eigenstates can be measured using several different analysis strategies. Perhaps the most elegant but also the most complex is time-dependent angular analysis of the $B_s^0 \to J/\psi\phi$ mode, where the flavour of the B_s^0 meson is identified using, for instance, opposite-side flavour-tagging techniques. The analysis is capable of extracting the average decay width Γ_s, decay

width difference $\Delta\Gamma_s$ the weak phase ϕ_s and decay amplitudes $|A_i|^2$ of the $B_s^0 \to J/\psi\phi$ mode simultaneously along with a set of other parameters related to B_s^0 decay.

An alternative, simplified method to flavour-tagged time-dependent angular analysis is to measure the effective lifetime [5] of the B_s^0 meson decaying in the $J/\psi\phi$ state. The effective lifetime measurement does not require tagging or angular analysis, unlike the ϕ_s study. The effective lifetime obtained from the analysis is proportional to all key results Γ_s, $\Delta\Gamma_s$, ϕ_s and $|A_i|^2$ of the flavour-tagged time-dependent angular analysis. Hence, the method is complementary to full angular analysis and can be used to independently cross-check the results given by the more advanced measurements.

Effective lifetime measurements have been previously done by the LHCb experiment [6] at LHC and the CDF [7] and D0 [8] collaborations at the Tevatron. Many flavour tagging studies and weak phase analyses have been made in the ATLAS [9, 10], the CMS [11, 12] and the LHCb experiments [13, 14], and earlier by the CDF [15, 16] and the D0 [17, 18] collaborations.

The ATLAS and the CMS detectors are designed to study a wide range of particle physics phenomena from energies of a few hundred MeV to about 10 TeV. The LHCb apparatus is devoted to studying the physics of b quarks and CP violation at energies ranging from tens of MeVs to a hundred GeV. The CMS apparatus has not been designed solely for B physics purposes, but it has a high-precision tracker and muon chambers to measure the momenta of charged kaons and pions as well as muons originating from J/ψ mesons. An additional benefit compared to the LHCb experiment or B factories is the higher luminosity delivered to the CMS detector. The CMS experiment has an efficient triggering system and thus can record substantial amounts of B meson decays. This allows for competitive, high-quality measurements in the field of B physics.

In this thesis, I present the analysis of the effective lifetime of the B_s^0 meson

decaying into the $J/\psi\phi$ final state. I also describe the design, validation, and performance of the two flavour tagging algorithms that identify the flavour of the B_s^0 mesons for the weak phase analysis. Moreover, the results of the flavour-tagged time-dependent angular analysis are briefly discussed. The measurements covered in the thesis are done using the data collected by the CMS detector in proton-proton collisions at a center-of-mass energy of 8 TeV. The data correspond to an integrated luminosity of 19.7 fb^{-1}.

2. Physics of neutral B mesons

2.1 Standard model of particle physics

The standard model of particle physics is a theory describing all the elementary particles discovered so far and including three of the four interactions between them. It is a renormalizible quantum field theory including special relativity, and combining the theory of the strong interaction, Quantum Chromo Dynamics [19, 20], as well as theory of the electroweak interaction [21–23].

The particles in the SM are represented as states created by quantum fields. The Lagrangian density of the SM is invariant under local SU(3) $\otimes$ SU(2) $\otimes$ U(1) gauge transformations. This invariance determines the interactions between the elementary particles: the SU(2) $\otimes$ U(1) group represents the electroweak interaction mediated by massive $W^{\pm}$ and Z^0 bosons and the massless photon γ. The SU(3) group describes the strong interaction mediated by eight massless gluons g. All the force carriers are spin-1 bosons.

The elementary particles of the SM are divided into two distinct groups of spin-1/2 fermions: the quarks and the leptons. Both fermion groups contain three families of particles organised in left-handed SU(2) doublets and right-handed singlets. Quarks are labelled as up-type (u,c,t) and down-type (d,s,b) species whilst leptons can be grouped into charged and massive particles (e, μ, τ) as well as neutral and nearly massless neutrinos (ν_e, ν_μ, ν_τ).

The masses of the $W^{\pm}$ and Z^0 bosons arise from a spontaneous breaking of SU(2) $\otimes$ U(1) symmetry called the Higgs mechanism [24–29]. This requires the last piece in the SM particle family: the spin-0 scalar boson. The spontaneous symmetry breaking is achieved by adding a scalar potential $V(\phi)$ in the SM Lagrangian: $V(\phi) = \mu^2\phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2$, where $\mu^2 < 0$ and $\lambda > 0$, and Φ is a doublet of two complex scalar fields. The Lagrangian with the scalar potential is invariant under SU(2) $\otimes$ U(1) transformations, but the ground state is not. The SU(2) $\otimes$ U(1) symmetry is broken into a U(1) symmetry - the $W^{\pm}$ and Z^0 bosons obtain their masses but photon stays massless: for each broken symmetry, a massive boson will appear.

The fermions interact with the scalar field Φ via the gauge-invariant Yukawa couplings. Diagonalizing the Yukawa couplings determines the mass eigenstates. The Cabibbo-Kobayashi-Maskawa (CKM) [30, 31] matrix transforms the mass eigenstates into the weak interaction eigenstates.

2.2 CKM matrix

Weak interactions mediated by the $W^{\pm}$ boson are the only interactions in the SM changing the quark flavour at the tree level. The part of the weak interaction Lagrangian is defined as

$$\mathcal{L} = -\frac{g}{2\sqrt{2}}\left(\bar{u}_i W_\mu^+ \gamma^\mu(1-\gamma^5)d_i' + \bar{d}_i' W_\mu^- \gamma^\mu(1-\gamma^5)u_i\right), \tag{2.1}$$

where u_i is an up-type quark field in the mass eigenstate, d_i' is a down-type quark field in the weak interaction basis and $W_\mu\gamma^\mu(1-\gamma^5)$ is the left-handed operator which connects the two bases.

The weak eigenstates are obtained by rotating the mass eigenstates by the

CKM matrix $V_{\rm CKM}$ [30, 31]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\rm CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.2}$$

The matrix element V_{ij} couples the up-type quark i to the down-type quark j, and $|V_{ij}|^2$ gives the transition probability between the quark species $i \leftrightarrow j$.

The 18 parameters of the CKM matrix are reduced to nine, by the unitarity condition $V_{\rm CKM}^\dagger V_{\rm CKM} = 1$. Five more parameters can be absorbed into unobservable quark phases. The matrix thus has four free parameters: one complex phase δ, and three mixing angles θ_{ij} [32].

The CKM matrix can be represented as a product of three rotation matrices comprising the mixing angles, $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$, and the complex phase δ.

$$\begin{aligned} V_{\rm CKM} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{12} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{2.3}$$

The relative size of the matrix elements are seen in Wolfenstein parametrisation [33], which has the real valued parameters A, λ, ρ and η given by [2]

$$\begin{aligned} s_{12} &= \lambda = \frac{|V_{us}|}{\sqrt{|V_{us}|^2 + |V_{ud}|^2}} \\ s_{23} &= A\lambda^2 = \lambda \left|\frac{V_{cb}}{V_{us}}\right| \\ s_{13}e^{i\delta} &= A\lambda^3(\rho + i\eta) = V_{us}^*, \end{aligned}$$

where the measured values of the parameters λ and A are $\lambda = 0.2256 \pm 0.0050$ and $A = 0.811 \pm 0.026$. The parameters ρ and η are often expressed with experimental

variables $\bar{\rho}$ and $\bar{\eta}$. They are related to parameters ρ and η and the CKM matrix elements with

$$A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho}+i\bar{\eta})\sqrt{1-A^2\lambda^4}}{\sqrt{1-\lambda^2}[1-A^2\lambda^4(\bar{\rho}+i\bar{\eta})]}$$
$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}.$$

The measured values of the variables are $\bar{\rho} = 0.124^{+0.019}_{-0.018}$ and $\bar{\eta} = 0.356 \pm 0.011$ [2].

The CKM matrix up $O(\lambda^5)$ is defined as

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda - A^2\lambda^5(\rho + i\eta - \frac{1}{2}) & 1 - \frac{\lambda^2}{2} - \lambda^4(\frac{1}{8} - \frac{A}{2}) & A\lambda^2 \\ A\lambda^3[1 - (\rho + i\eta)(1 - \frac{\lambda^2}{2})] & -A\lambda^2 - A\lambda^4(\rho + i\eta - \frac{1}{2}) & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix}. \tag{2.4}$$

By comparing the matrices in (2.2) and (2.4), the diagonal elements V_{ud}, V_{cs} and V_{tb} are observed to be the largest and near unity. These elements refer to the transitions between the quarks within the same family. The other elements describing the cross-generation transitions are substantially smaller: V_{us} and V_{cd} are proportional to λ, V_{ts} and V_{cb} to λ^2 and V_{td} and V_{ub} to λ^3. The parameter η in the Wolfenstein representation causes the CP violation in the standard model. CP violation is suppressed in the SM, since η is multiplied by a small factor λ^3 [32].

2.2.1 Unitarity triangles

The unitarity condition of the CKM matrix gives three conditions to the off-diagonal matrix elements,

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{2.5}$$
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.6}$$
$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{2.7}$$

Three complex numbers, which add up to zero, can be visualized as the sides of a triangle. It is customary to normalize one side of each triangle to unity leading to

$$\frac{V_{ud}V_{us}^*}{V_{cd}V_{cs}^*} + \frac{V_{cd}V_{cs}^*}{V_{cd}V_{cs}^*} + \frac{V_{td}V_{ts}^*}{V_{cd}V_{cs}^*} = 0 \tag{2.8}$$

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0. \tag{2.9}$$

$$\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*} + \frac{V_{cs}V_{cb}^*}{V_{cs}V_{cb}^*} + \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} = 0 \tag{2.10}$$

A visual interpretation of Equations (2.9) - (2.10) is shown in Figure 2.1. Each angle represents a transition of a b quark to another quark: b $\rightarrow$ (u,c,t). The angles of the triangles on the right-hand side can be accessed with decays of B_d^0 mesons whereas the angles of the triangle in the left-hand side can be measured with the decays of B_s^0 mesons. Since the angle β_s is small, the triangle related to Equation (2.10) is nearly flat.

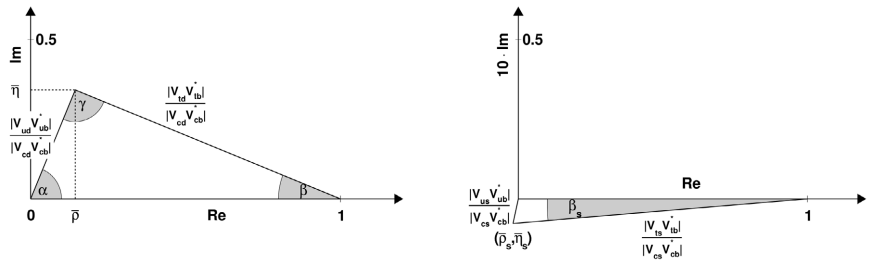


Figure 2.1: Unitarity triangles whose angles contain b $\rightarrow$ (u,c,t) transitions. The triangle on the left corresponds to Equation (2.9), and the triangle on the right refers to Equation (2.10). Both images of the triangles are taken from [34].

The angles of the left triangle in Figure 2.1 are defined as

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{2.11}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{2.12}$$

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right). \tag{2.13}$$

The angle β_s of the right-hand triangle in Figure 2.1 is given by

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{2.14}$$

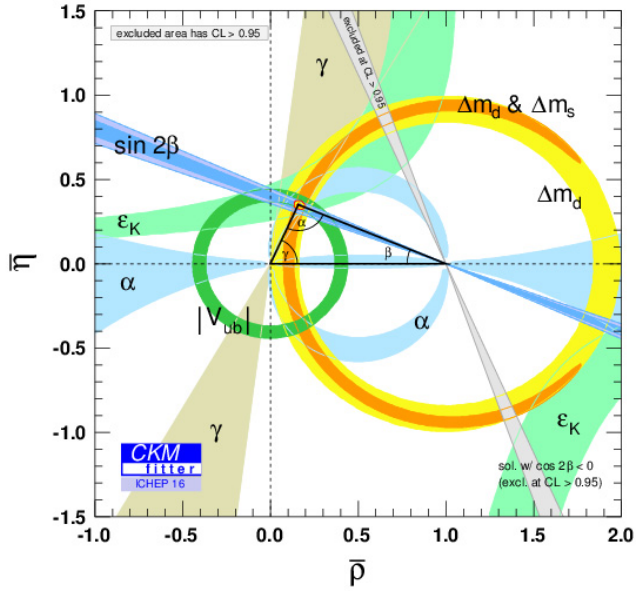
This angle is proportional to the weak mixing phase ϕ_s of the B_s^0 meson system via $\phi_s \approx -2\beta_s = -0.0370 \pm 0.0006$ rad [4], and will be discussed in more detail in Section 2.3.6.

Measuring the magnitudes of the CKM matrix elements and the angles of the unitarity triangles is one of the main goals of flavour physics experiments. The measurements test the unitarity of the CKM matrix. Measurements of the CKM parameters using the decays including the mixing of the neutral mesons are especially interesting, since they can be used to extract the CKM elements only if the SM is assumed. Thus, these measurements are particularly sensitive to contributions of new physics beyond the standard model.

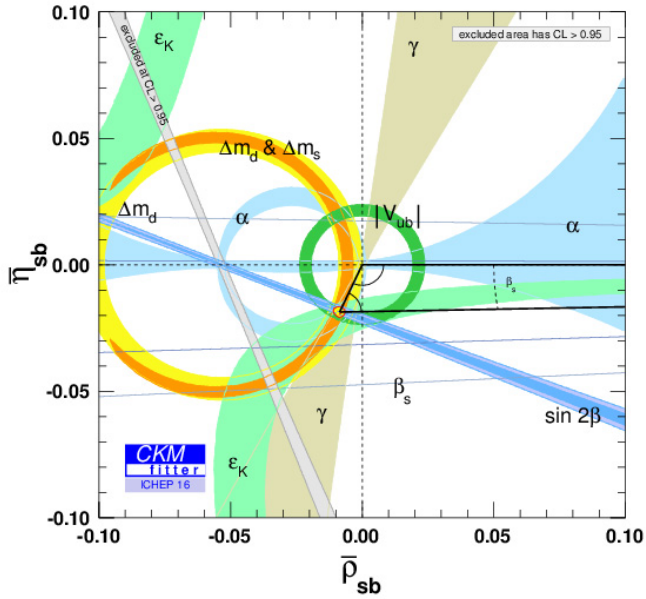
The validity of the SM in the flavour sector is tested using global fits of a large number of experimental results, since the CKM matrix elements are free parameters of the SM. Two such fits [35] showing the unitary triangles from Equations (2.9) - (2.10) in the $(\bar{\rho}, \bar{\eta})$ plane are presented in Figure 2.2. The allowed apex regions of the angles α and β_s are indicated with a small yellow dot. The dot corresponds to a 68% confidence level. A good agreement with the unitarity conditions is obtained.

2.3 Neutral B meson systems

In the next few sections a short overview of the B meson phenomenology is presented. Good references regarding CP violation and the physics of neutral B mesons are for example [2, 36, 37], whose notations the text in Sections 2.3.1 - 2.3.6 also follows. References related to the weak mixing phase of Section 2.3.6 can be found in [38, 39]. The description of the $B_s^0 \to J/\psi\phi$ decay rate in the transversity basis is covered in Section 2.4 and based on [40]. The formalism for the effective lifetime can be found



(a)



(b)

Figure 2.2: Global fits to data related to b quark transitions. A good agreement with the unitarity conditions is obtained. The plots are taken from [35].

in [5]. The description in Section 2.4.2 - 2.4.3 of flavour tagging techniques and the raw B_d^0 mixing asymmetry are based on [2, 38, 41, 42].

2.3.1 Time evolution of the neutral B meson states

The superposition of a neutral B_q^0 meson of flavour $(\bar{b}q)$ and its antiparticle $\bar{B}_q^0$ of flavour $(b\bar{q})$ may be expressed as

$$|\psi(t)\rangle = a(t)|B_q^0\rangle + b(t)|\bar{B}_q^0\rangle. \tag{2.15}$$

Denoting $|B_q^0\rangle = \binom{1}{0}$ and $|\bar{B}_q^0\rangle = \binom{0}{1}$, time dependence of the coefficients is determined by the Schrödinger equation

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathbf{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{2.16}$$

where $\mathbf{H}$ is an effective Hamiltonian. The Hamiltonian is given by hermitean mass and decay matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ defined as

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M - \frac{i}{2}\Gamma \end{pmatrix} = \begin{pmatrix} H_0 & H_{12} \\ H_{21} & H_0 \end{pmatrix}. \tag{2.17}$$

The diagonal elements of $\mathbf{M}$ and $\mathbf{\Gamma}$ can be denoted as $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$ due to the conservation of CPT symmetry. The off-diagonal elements of the mass matrix determine the mixing of the neutral mesons $B \leftrightarrow \bar{B}$. A more detailed description of the matrix elements is provided in Section 2.3.2. The eigenvalues of the Hamiltonian are

$$\lambda_H = H_0 + \sqrt{H_{12}H_{21}} = M - \frac{i}{2}\Gamma + \sqrt{(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)(M_{12} - \frac{i}{2}\Gamma_{12})}. \tag{2.18}$$

$$\lambda_L = H_0 - \sqrt{H_{12}H_{21}} = M - \frac{i}{2}\Gamma - \sqrt{(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)(M_{12} - \frac{i}{2}\Gamma_{12})}. \tag{2.19}$$

The eigenvalues correspond to the eigenvectors

$$v_H = \frac{1}{\sqrt{|p|^2 + |q|^2}}\begin{pmatrix} p \\ -q \end{pmatrix}, \quad v_L = \frac{1}{\sqrt{|p|^2 + |q|^2}}\begin{pmatrix} p \\ +q \end{pmatrix}, \tag{2.20}$$

where the ratio of the coefficients q, p fulfils the condition

$$\frac{q}{p} = -\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.21}$$

The eigenvectors can be expressed in the basis of the flavour states as

$$v_H = |\mathrm{B_H}\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}\left(p|\mathrm{B_q^0}\rangle - q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle\right) \tag{2.22}$$

$$v_L = |\mathrm{B_L}\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}\left(p|\mathrm{B_q^0}\rangle + q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle\right), \tag{2.23}$$

which are called the heavy and light B meson eigenstates. The time evolution of the general state in Equation (2.15) can now be written in the basis of the eigenstates $|\mathrm{B_L}\rangle$ and $|\mathrm{B_H}\rangle$ as

$$|\psi(t)\rangle = c_H e^{-i\lambda_H t}|\mathrm{B_H}\rangle + c_L e^{-i\lambda_L t}|\mathrm{B_L}\rangle. \tag{2.24}$$

The coefficients c_L, c_H are determined by the initial conditions. For example, if the time evolution of the B meson starts from one of the flavour states, then $|\psi(0)\rangle = |\mathrm{B_q^0}\rangle$ and $|\psi(0)\rangle = |\bar{\mathrm{B}}_\mathrm{q}^0\rangle$. The time evolution can be written as

$$|\mathrm{B_q^0}(t)\rangle = g_+(t)|\mathrm{B_q^0}\rangle + \tfrac{q}{p}g_-(t)|\bar{\mathrm{B}}_\mathrm{q}^0\rangle, \tag{2.25}$$

$$|\bar{\mathrm{B}}_\mathrm{q}^0(t)\rangle = g_-(t)\tfrac{p}{q}|\mathrm{B_q^0}\rangle + g_+(t)|\bar{\mathrm{B}}_\mathrm{q}^0\rangle, \tag{2.26}$$

where the time evolution of the $\mathrm{B_q^0}$ ($\bar{\mathrm{B}}_\mathrm{q}^0$) flavour state assumes that the meson was produced as $\mathrm{B_q^0}$ ($\bar{\mathrm{B}}_\mathrm{q}^0$). The functions $g_\pm(t)$ are

$$g_\pm(t) = \frac{1}{2}\left(e^{-i\lambda_L t} \pm e^{-i\lambda_H t}\right) = \frac{1}{2}\left(e^{-i\left(m_L - \frac{i}{2}\Gamma_L\right)t} \pm e^{-i\left(m_H - \frac{i}{2}\Gamma_H\right)t}\right), \tag{2.27}$$

where $\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L}$ are the eigenvalues of the Hamiltonian.

2.3.2 Matrix elements of the Hamiltonian

The matrix elements of the Hamiltonian describe the mixing and decay processes of the neutral B mesons. M_{12} determines the transition $\bar{\mathrm{B}}_\mathrm{q}^0 \to \mathrm{B_q^0}$ and is given by

$$M_{12} = \frac{G_F^2 m_W^2}{12\pi^2} f_{B_q}^2 m_{B_q} B_{B_q} F^*, \tag{2.28}$$

where the factor F reads

$$\begin{aligned} F &= \eta_1 (V_{cq}V_{cb}^*)^2 S_0(x_c) + \eta_2 (V_{tq}V_{tb}^*)^2 S_0(x_t) + 2\eta_3 (V_{cq}V_{cb}^*)^2 (V_{tq}V_{tb}^*)^2 S_0(x_c, x_t) \\ &\approx \eta_2 (V_{tq}V_{tb}^*)^2 S_0(x_t). \end{aligned} \tag{2.29}$$

The function $S_0(x_q)$ is a hadronic form factor defined as

$$S_0(x_q) = \frac{x_q}{(1-x_q^2)} \left\{ 1 - \frac{11x_q}{4} + \frac{x_q^2}{4} - \frac{3x^2 \ln(x_q)}{2(1-x_q)} \right\}, \tag{2.30}$$

where $x_q = \frac{m_q^2}{m_W^2}$ and the coefficients η_i are QCD correction factors. More details about the function $S_0(x_c, x_t)$ is given in Appendix B of Reference [36]. Because $S_0(x_t) >> S_0(x_c), S_0(x_c, x_t)$, and since the products of the CKM matrix elements $(V_{cq}V_{cb}^*)^2$ and $(V_{tq}V_{tb}^*)^2$ have similar order of magnitudes for $\mathrm{B_d^0}$ and $\mathrm{B_s^0}$ mesons, the second term in Equation (2.29) dominates.

Remaining undefined quantities in Equation (2.29) are m_{B_q} that refers to the mass of the $\mathrm{B_q^0}$ meson and the bag parameter B_{B_q} that parametrises the hadronic matrix element of $\mathrm{B_q^0}$ mixing. The parameter f_{B_q} and G_F are the B meson decay constant and the Fermi constant. Using the above approximations and simplifications, matrix element M_{12} becomes

$$M_{12} \approx \frac{G_F^2 m_W^2}{12\pi^2} f_{B_q}^2 m_{B_q} B_{B_q} \eta_2 S_0(x_t) (V_{tb}V_{tq}^*)^2. \tag{2.31}$$

The matrix element Γ_{12} is a product of the decay amplitudes A_f and $\bar{A}_f$ of the $\mathrm{B_q^0}$ and $\bar{\mathrm{B}}_\mathrm{q}^0$ mesons, summed over the physical final states f accessible to $\mathrm{B_q^0}$ and $\bar{\mathrm{B}}_\mathrm{q}^0$,

$$\Gamma_{12} = \textstyle\sum_f A_f^* \bar{A}_f \tag{2.32}$$

$$A_f = \langle f | \mathbf{H} | \mathrm{B_q^0} \rangle \tag{2.33}$$

$$\bar{A}_f = \langle f | \mathbf{H} | \bar{\mathrm{B}}_\mathrm{q}^0 \rangle. \tag{2.34}$$

One obtains [2, 43]

$$\begin{aligned} \Gamma_{12} = \frac{G_F^2 m_W^2}{8\pi} m_b^2 \eta_B B_{B_q} f_{B_q}^2 \{ &(V_{tq}^* V_{tb})^2 \\ + V_{tq}^* V_{tb} V_{cq}^* V_{cb} O\left(\frac{m_c^2}{m_b^2}\right) &+ (V_{cq}^* V_{cb})^2 O\left(\frac{m_c^4}{m_b^4}\right) \}. \end{aligned} \tag{2.35}$$

The ratio of the magnitudes of the decay and mass matrix elements is small,

$$|\frac{\Gamma_{12}}{M_{12}}| \approx \frac{3\pi m_b^2}{2m_W^2 S_0(x_t)} \approx \frac{m_b^2}{m_t^2} \approx 10^{-3}. \tag{2.36}$$

The diagonal matrix elements $M_{11} = M_{22} = M$ depend on processes where the neutral meson oscillates back to its initial flavour state before it decays. The element M_{11} refers to a case where the meson is created and decays as $\mathrm{B_q^0}$ and M_{22} to case where the meson is born and decays as $\bar{\mathrm{B}}_\mathrm{q}^0$. The matrix elements $\Gamma_{11} = \Gamma_{22} = \Gamma$ describe decay amplitudes $\mathrm{B_q^0} \to f$ and $\bar{\mathrm{B}}_\mathrm{q}^0 \to f$ determined as

$$\Gamma_{11} = \textstyle\sum_f |A_f|^2 \tag{2.37}$$

$$\Gamma_{22} = \textstyle\sum_f |\bar{A}_f|^2. \tag{2.38}$$

2.3.3 Decay widths and masses

The physical B meson states, $|\mathrm{B_H}\rangle$ and $|\mathrm{B_L}\rangle$, have distinct masses $m_{H,L}$ and decay widths $\Gamma_{H,L}$, that are the eigenvalues of the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$. The masses and the decay widths can be defined with the eigenvalues of the Hamiltonian in Equation (2.17),

$$\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L}, \tag{2.39}$$

and particularly with

$$m_{H,L} = \mathrm{Re}(\lambda_{H,L}) \tag{2.40}$$

$$\Gamma_{H,L} = -2\mathrm{Im}(\lambda_{H,L}). \tag{2.41}$$

Furthermore, the mass and width differences Δm and $\Delta\Gamma$ can be presented as

$$\Delta m = m_H - m_L = 2\mathrm{Re}\left(\sqrt{H_{12}H_{21}}\right) \tag{2.42}$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H = -4\mathrm{Im}\left(\sqrt{H_{12}H_{21}}\right). \tag{2.43}$$

Using the definitions in Equations (2.42) - (2.43) and transforming the complex number $\sqrt{H_{12}H_{21}}$ of Equation (2.17) into the polar form, it is possible to show that

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2 \tag{2.44}$$

$$\Delta m \Delta\Gamma = -4|M_{12}||\Gamma_{12}|\cos(\phi_\Gamma - \phi_M), \tag{2.45}$$

where $\phi_M = \arg(M_{12})$ and $\phi_\Gamma = \arg(\Gamma_{12})$. The measurements show that $\Delta m >> \Delta\Gamma$ for both neutral B meson systems (c.f. Table 2.1) and that $|M_{12}| >> |\Gamma_{12}|$ (c.f. Equation (2.36)). Hence, the mass and decay widths can be approximated as

$$\Delta m = 2|M_{12}| \tag{2.46}$$

$$\Delta\Gamma = -2|\Gamma_{12}|\cos(\phi_\Gamma - \phi_M). \tag{2.47}$$

The same approximations allow for writing the ratio of the mixing parameters q, p as

$$\frac{q}{p} = -\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \approx -\sqrt{\frac{M_{12}^*}{M_{12}}} = -e^{-i\phi_M}. \tag{2.48}$$

To simplify the decay rate formulas of the neutral B mesons to be discussed later in this chapter, it is convenient to define the average decay width given as

$$\Gamma = \frac{1}{2}(\Gamma_L + \Gamma_H). \tag{2.49}$$

The average decay width as well as the mass and decay width differences for both B_s^0 and B_d^0 mesons based on several measurements are shown in Table 2.1. The mass difference determines the mixing frequency within the flavour states $B_q^0 \leftrightarrow \bar{B}_q^0$. The B_s^0 meson has a much larger mixing frequency than the B_d^0 hadron and oscillates within its flavour states several times during its average lifetime, whereas B_d^0 does this only once or twice.

The decay width difference of the physical states determines the difference of the lifetimes $\tau_{H,L}$ of these states with $\tau_{H,L} = 1/\Gamma_{H,L}$. For B_s^0, the lifetimes of the physical states are very different, as indicated by large relative decay width difference $|\Delta\Gamma|/\Gamma \approx 12.4$ % [2]. The B_d^0 system, on the other hand, has a negligible width difference and its physical states have practically equal lifetimes.

	B_s^0	B_d^0
Γ (ps^{-1})	0.6623 ± 0.0022	0.6579 ± 0.0017
$\lvert\Delta\Gamma\rvert/\Gamma$ (%)	12.4 ± 1.1	0.3 ± 1.5
Δm (ps^{-1})	17.757 ± 0.021	0.5096 ± 0.0034

Table 2.1: Average decay width, relative decay width difference and mass difference of the neutral B mesons as reported by the Particle Data Group [2].

2.3.4 Mass and CP eigenstates

The B_q^0 meson transforms under parity and charge conjugation as

$$\hat{P}|B_q^0\rangle = -|B_q^0\rangle \tag{2.50}$$

$$\hat{C}|B_q^0\rangle = +|\bar{B}_q^0\rangle. \tag{2.51}$$

Hence, the flavour states transform under $\hat{C}\hat{P}$ as

$$\hat{C}\hat{P}|B_q^0\rangle = -\hat{C}|B_q^0\rangle = -|\bar{B}_q^0\rangle \tag{2.52}$$

$$\hat{C}\hat{P}|\bar{B}_q^0\rangle = -\hat{C}|\bar{B}_q^0\rangle = -|B_q^0\rangle. \tag{2.53}$$

The CP eigenstates $|B_q^{even}\rangle$ and $|B_q^{odd}\rangle$ are

$$|B_q^{even}\rangle = \frac{1}{\sqrt{2}}\left(|B_q^0\rangle - |\bar{B}_q^0\rangle\right) \tag{2.54}$$

$$|B_q^{odd}\rangle = \frac{1}{\sqrt{2}}\left(|B_q^0\rangle + |\bar{B}_q^0\rangle\right), \tag{2.55}$$

where the CP-even state has an eigenvalue of $\eta_{CP} = +1$ and the CP-odd eigenstate has $\eta_{CP} = -1$.

Since the mixing phase ϕ_M in Equation (2.48) is predicted and indirectly measured to be small, $\frac{q}{p} \approx -1$. Using this approximation and the resulting fact that $|\frac{q}{p}| = 1$, we can write the mass eigenstates in Equations (2.22) and (2.23) in the new

form

$$|\mathrm{B_L}\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}\left(p|\mathrm{B_q^0}\rangle + q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle\right) \approx e^{i\theta_{q,p}}|\mathrm{B_q^{even}}\rangle \tag{2.56}$$

$$|\mathrm{B_H}\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}\left(p|\mathrm{B_q^0}\rangle - q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle\right) \approx e^{i\theta_{q,p}}|\mathrm{B_q^{odd}}\rangle, \tag{2.57}$$

where $\theta_{q,p}$ is phase of the complex parameters q and p. From Equations (2.56) and (2.57), it is clear that the heavy mass eigenstate corresponds to the CP-odd state and light eigenstate to the CP-even state. This assumption is helpful when we discuss the $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ decay rate.

2.3.5 Decay rates and mixing probabilities

Before determining the time dependent decay rates and mixing probabilities, it is useful to note some identities for the functions $g_\pm(t)$

$$\begin{aligned}
|g_+(t)|^2 &= \tfrac{1}{2}e^{-\Gamma t}\left(\cosh\left(\tfrac{\Delta\Gamma}{2}t\right) + \cos(\Delta m t)\right)\\
|g_-(t)|^2 &= \tfrac{1}{2}e^{-\Gamma t}\left(\cosh\left(\tfrac{\Delta\Gamma}{2}t\right) - \cos(\Delta m t)\right)\\
g_-^*(t)g_+(t) &= \tfrac{1}{2}e^{-\Gamma t}\left(-\sinh\left(\tfrac{\Delta\Gamma}{2}t\right) - i\sin(\Delta m t)\right)\\
g_+^*(t)g_-(t) &= \tfrac{1}{2}e^{-\Gamma t}\left(-\sinh\left(\tfrac{\Delta\Gamma}{2}t\right) + i\sin(\Delta m t)\right).
\end{aligned} \tag{2.58}$$

The mixing probability for a neutral B meson is

$$P_{mix}^{\mathrm{B_q^0}}(t) = |\langle\bar{\mathrm{B}}_\mathrm{q}^0|\mathrm{B_q^0}(t)\rangle|^2 = |g_-(t)\frac{q}{p}|^2 = \frac{1}{2}|\frac{q}{p}|^2 e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \cos(\Delta m t)\right) \tag{2.59}$$

$$P_{mix}^{\bar{\mathrm{B}}_\mathrm{q}^0}(t) = |\langle\mathrm{B_q^0}|\bar{\mathrm{B}}_\mathrm{q}^0(t)\rangle|^2 = |g_-(t)\frac{p}{q}|^2 = \frac{1}{2}|\frac{p}{q}|^2 e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \cos(\Delta m t)\right) \tag{2.60}$$

where the former and latter equations describe the mixing probability of a B meson initially produced as $\mathrm{B_q^0}$ and $\bar{\mathrm{B}}_\mathrm{q}^0$, respectively. The probability that the meson remains unmixed during its lifetime is written as

$$P_{unmix}^{\mathrm{B_q^0}}(t) = P_{unmix}^{\bar{\mathrm{B}}_\mathrm{q}^0}(t) = |g_+(t)|^2 = \frac{1}{2}e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \cos(\Delta m t)\right). \tag{2.61}$$

It is helpful to simplify the expressions by introducing a parameter λ_f that is closely related to mixing induced CP violation

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \tag{2.62}$$

where A_f and $\bar{A}_f$ are the decay amplitudes of B^0_q and $\bar{\mathrm{B}}^0_\mathrm{q}$ mesons shown in Equations (2.34) - (2.34). Using the parameter λ_f and identities (2.58) for $g_\pm(t)$ the time dependent decay rate to the final state f for a neutral B meson initially produced as B^0_q is

$$\begin{aligned}
\frac{d\Gamma\left(\mathrm{B}^0_\mathrm{q} \to f\right)}{dtN_f} &= |\langle f|\mathrm{B}^0_\mathrm{q}(t)\rangle|^2 = |g_+(t)A_f + \frac{q}{p}g_-(t)\bar{A}_f|^2 \\
&= |A_f|^2\{|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + \lambda_f^* g_+(t)g_-^*(t) + \lambda_f g_-(t)g_+^*(t)\} \\
&= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\{\left(1+|\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \left(1-|\lambda_f|^2\right)\cos\left(\Delta m t\right) \\
&- 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\mathrm{Re}(\lambda_f) - 2\sin\left(\Delta m t\right)\mathrm{Im}(\lambda_f)\}
\end{aligned} \tag{2.63}$$

The differential decay rate of a B meson originally produced as $\bar{\mathrm{B}}^0_\mathrm{q}$ becomes

$$\begin{aligned}
\frac{d\Gamma\left(\bar{\mathrm{B}}^0_\mathrm{q} \to f\right)}{dtN_f} &= |\langle f|\bar{\mathrm{B}}^0_\mathrm{q}(t)\rangle|^2 = |g_+(t)\bar{A}_f + \frac{p}{q}g_-(t)A_f|^2 \\
&= |A_f|^2|\frac{p}{q}|^2\{|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + \lambda_f^* g_-(t)g_+^*(t) + \lambda_f g_+(t)g_-^*(t)\} \\
&= \frac{1}{2}|A_f|^2|\frac{p}{q}|^2 e^{-\Gamma t}\{\left(1+|\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \left(1-|\lambda_f|^2\right)\cos\left(\Delta m t\right) \\
&- 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\mathrm{Re}(\lambda_f) + 2\sin\left(\Delta m t\right)\mathrm{Im}(\lambda_f)\},
\end{aligned} \tag{2.64}$$

where the factor N_f is a normalization constant. The decay rates to the CP conjugated final state $\bar{f}$ are obtained by substituting $A_f \to A_{\bar{f}}$, $\bar{A}_f \to \bar{A}_{\bar{f}}$ and $\lambda_f \to \lambda_{\bar{f}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}}$.

2.3.6 CP violation in the meson decays

CP violation can manifest itself in three different ways:

- CP violation in decay, which is also called direct CP violation
- CP violation in mixing called indirect CP violation
- CP violation in the interference between mixing and decay, which is called mixing induced CP violation

CP violation in decay

Direct CP violation occurs when the decay rate of a meson M into a final state f differs from the decay rate of the anti-meson $\overline{M}$ into the CP-conjugated state $\bar{f}$. In terms of the decay amplitudes A_f and $\bar{A}_{\bar{f}}$ this implies

$$|\frac{\bar{A}_{\bar{f}}}{A_f}| \neq 1. \tag{2.65}$$

CP violation in decay is possible for any hadrons. It requires that the total decay amplitude is the sum of at least two terms with distinct phases,

$$A_f = \sum_{i=1}^{n} |A_i| e^{i(\delta_i + \phi_i)} \tag{2.66}$$

$$\bar{A}_{\bar{f}} = \sum_{i=1}^{n} |A_i| e^{i(\delta_i - \phi_i)}, \tag{2.67}$$

where δ_i and ϕ_i are strong and weak interaction phases. This enables a relative phase difference between the different decay amplitudes and leads to Equation (2.65).

CP violation in mixing

Indirect CP violation occurs for neutral B mesons when the parameters q and p in Equations (2.22) - (2.23) have different magnitudes

$$|\frac{q}{p}| \neq 1. \tag{2.68}$$

In practice, this means that the probability for a neutral meson B to oscillate to its antiparticle $\bar{\mathrm{B}}$ is different than the probability that $\bar{\mathrm{B}}$ mixes with B, that is $P_{mix}^{\mathrm{B_q^0}} \neq P_{mix}^{\bar{\mathrm{B}}_\mathrm{q}^0}$. From Equations (2.59)-(2.60) it is easily seen that indeed the mixing probabilities are different only if $|\frac{q}{p}| \neq 1$ holds.

CP violation in interference between mixing and decay

CP violation in interference between mixing and decay can occur when a neutral $\mathrm{B_q^0}$ meson and its antiparticle $\bar{\mathrm{B}}_\mathrm{q}^0$ can decay into the same final state f e.g. as $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$. The decay can then occur with or without mixing: $\mathrm{B_q^0} \to f$ or $\mathrm{B_q^0} \to \bar{\mathrm{B}}_\mathrm{q}^0 \to f$.

For this type of CP violation, the quantity of interest is the parameter λ_f defined by (2.62). CP violation occurs when λ_f is complex,

$$\mathrm{Im}(\lambda_f) \neq 0, \tag{2.69}$$

implying a non-trivial phase $\phi_f = \arg(\lambda_f)$. An important feature of CP violation in interference between mixing and decay is that it can occur even if there is neither direct CP nor indirect CP violation. In this case

$$\arg(\lambda_f) + \arg(\lambda_{\bar{f}}) \neq 0. \tag{2.70}$$

Weak mixing phase in $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ decays

The weak mixing phase ϕ_s in $\bar{\mathrm{b}} \to \mathrm{c\bar{c}\bar{s}}$ transitions arises from the phase mismatch between the amplitudes of the direct decay and the decay involving mixing. Experimentally this phase can be measured in $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ decay. The Feynman diagrams of $\mathrm{B_s^0}$ mixing is illustrated in Figure 2.3, and the diagrams of a direct decay and decay with a gluonic transition are presented in Figure 2.4. The total $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ decay amplitude is the sum of the amplitudes from the tree-level decay amplitude

(T) and penguin contributions ($P_{u,c,t}$). The total decay amplitude is thus

$$\begin{aligned} A_{J\psi\phi} &= V_{cs}V_{cb}^{*}T + V_{us}V_{ub}^{*}P_u + V_{cs}V_{cb}^{*}P_c + V_{ts}V_{tb}^{*}P_t \\ &= V_{cs}V_{cb}^{*}\left(T + P_c - P_t\right) + V_{us}V_{ub}^{*}(P_u - P_t) \approx V_{cs}V_{cb}^{*}\left(T + P_c - P_t\right). \end{aligned} \tag{2.71}$$

The second equality follows from the unitarity condition (2.7), $V_{ts}V_{tb}^{*} = -(V_{us}V_{ub}^{*} + V_{cs}V_{cb}^{*})$. The product product $V_{us}V_{ub}^{*}$ is proportional to $A\lambda^4$ and $V_{cs}V_{cb}^{*} \propto A\lambda^2$ and $\lambda \approx 0.23$. Using this information and assuming that the tree level decay dominates the penguin contributions the term $V_{us}V_{ub}^{*}$ can be neglected from (2.71)

Since J/$\psi\phi$ is a CP eigenstate, $|\overline{\mathrm{J}/\psi\phi}\rangle = CP|\mathrm{J}/\psi\phi\rangle = n_{\mathrm{J}/\psi\phi}|\mathrm{J}/\psi\phi\rangle$, it follows that $\bar{A}_{J/\psi\phi} = n_{J/\psi\phi}\bar{A}_{\overline{J/\psi\phi}}$. Using Equations (2.66) and (2.67), decay amplitudes $A_{J/\psi\phi}$ and $\bar{A}_{J/\psi\phi}$ can be written as

$$A_{J/\psi\phi} = |V_{cs}V_{cb}^{*}|e^{-i\phi_D}K \tag{2.72}$$

$$\bar{A}_{J/\psi\phi} = n_{J/\psi\phi}\bar{A}_{\overline{J/\psi\phi}} = n_{J/\psi\phi}|V_{cs}V_{cb}^{*}|e^{i\phi_D}K, \tag{2.73}$$

where $K = \{|T|e^{i\delta_T} + |P_c|e^{i\delta_c} + |P_t|e^{i\delta_t}\}$, δ_T, δ_c and δ_t are phases due to the strong interactions and the phase $\phi_D = \arg(V_{cs}^{*}V_{cb})$. The ratio of the amplitudes then simplifies to

$$\frac{\bar{A}_{J\psi\phi}}{A_{J\psi\phi}} = n_{J\psi\phi}\frac{V_{cs}^{*}V_{cb}}{V_{cs}V_{cb}^{*}} = n_{J\psi\phi}e^{2i\phi_D}. \tag{2.74}$$

Furthermore, using the approximation of Equation (2.48) and the expression for the matrix element M_{12} in Equation (2.31), q/p is given by

$$\frac{q}{p} \approx -\sqrt{\frac{M_{12}^{*}}{M_{12}}} = -\frac{V_{tb}^{*}V_{ts}}{V_{tb}V_{ts}^{*}} = -e^{-i\phi_M}, \tag{2.75}$$

where the phase $\phi_M = \arg(M_{12}) = 2\arg(V_{tb}V_{ts}^{*})$. Using Equations (2.74) and (2.75), the CP-violation parameter $\lambda_{J/\psi\phi}$ becomes

$$\lambda_{J/\psi\phi} = n_{J/\psi\psi}\frac{q}{p}\frac{\bar{A}_{J\psi\phi}}{A_{J\psi\phi}} = -n_{J/\psi\psi}e^{-i(\phi_M - 2\phi_D)} \equiv -n_{J/\psi\psi}e^{-i\phi_s}, \tag{2.76}$$

where $\phi_s \equiv (\phi_M - 2\phi_D)$ is the weak mixing phase of the $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ decay. $\lambda_{J/\psi\phi}$ can also be expressed in terms of the CKM matrix elements,

$$\lambda_{\mathrm{J}/\psi\phi} = -n_{\mathrm{J}/\psi\phi}\frac{V_{tb}^{*}V_{ts}}{V_{tb}V_{ts}^{*}}\frac{V_{cs}^{*}V_{cb}}{V_{cs}V_{cb}^{*}} = -n_{\mathrm{J}/\psi\phi}e^{2i\beta_s}, \tag{2.77}$$

where $\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right)$ is an angle of the unitarity triangle shown in Figure 2.1. In the standard model, $\phi_s = -2\beta_s$ when term $V_{us}V_{ub}^*$ arising from the loop diagrams in Equation (2.71) is neglected. The decay phase $\phi_D = \arg(V_{cs}^*V_{cb})$ is approximately zero, since $\mathrm{Im}(V_{cs}^*V_{cb})$ is order of $O(\lambda^7)$. Thus, the leading contribution to ϕ_s is due to the matrix element M_{12} of $\mathrm{B_s^0}$ - $\mathrm{\bar{B}_s^0}$ mixing. This is the reason for calling ϕ_s a weak *mixing* phase.

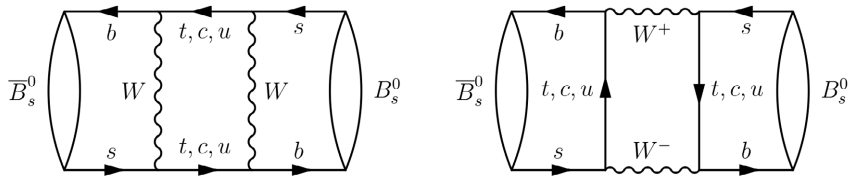


Figure 2.3: Feynman diagrams of the $\mathrm{B_s^0} \leftrightarrow \mathrm{\bar{B}_s^0}$ mixing taken from [44].

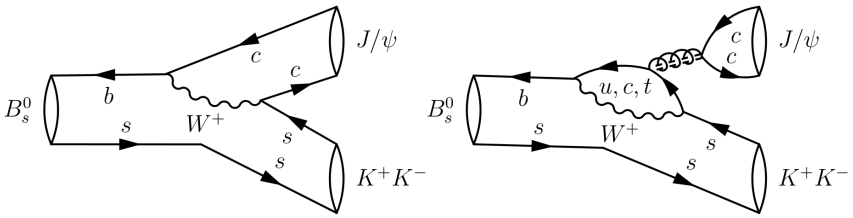


Figure 2.4: Decay of a $\mathrm{B_s^0}$ meson into a final state $\mathrm{J}/\psi\phi$ at tree-level (left) and through a penguin loop (right). The Feynman diagrams are taken from [44].

Effects of new physics on ϕ_s

New undiscovered particles could prominently reveal themselves in the $\mathrm{B_s^0}$ system by contributing to the mixing of the $\mathrm{B_s^0}$ mesons and thus shifting the value of ϕ_s from that predicted by the standard model. The matrix element M_{12} can be parametrised as

$$M_{12} = M_{12}^{SM}\Delta_s, \qquad \Delta_s = |\Delta_s|e^{i\phi_s^{\Delta}} \tag{2.78}$$

where M_{12}^{SM} is the standard model CKM matrix element and Δ_s is a new complex parameter that describes the beyond SM effects on the B_s^0 flavour oscillations. Parameter ϕ_s^Δ is the phase of Δ_s [35]. The weak mixing phase then becomes

$$\phi_s = \phi_s^{SM} + \phi_s^\Delta, \tag{2.79}$$

where $\phi_s^{SM} = -2\beta_s$ is the standard model contribution to the mixing phase. Δ_s would not only affect ϕ_s but also the mass and decay width differences of the B_s^0 meson [38]

$$\Delta m \ \to 2|M_{12}^{SM}||\Delta_s| \tag{2.80}$$

$$\Delta\Gamma \ \to -2|\Gamma_{12}^{SM}|\cos(\phi_\Gamma^{SM} - \phi_M^{SM} + \phi_s^\Delta). \tag{2.81}$$

Figure 2.5 shows the constraints on Δ_s from the data on B_s^0 mixing. The standard model prediction, $\mathrm{Re}(\Delta_s) = 1$ and $\mathrm{Im}(\Delta_s) = 0$ [35], is compatible with the measurements to 0.3σ. There is presently no evidence for new physics contributions to B_s^0 mixing.

2.4 Decay rate in the $B_s^0 \to J/\psi\phi$ channel

In the $B_s^0 \to J/\psi\phi$ decay a spin-0 B_s^0 meson decays to spin-1 J/ψ ($c\bar{c}$) and ϕ ($s\bar{s}$) mesons. Due to the conservation of total angular momentum $\vec{J}_{B_s} = \vec{L} + \vec{S}_{J/\psi} + \vec{S}_\phi$, the allowed angular momentum quantum numbers in the $J/\psi\phi$ final state are $L = 0, 1, 2$. Since both J/ψ and ϕ mesons have individual charge-conjugation eigenvalues $C_{J/\psi} = C_\phi = -1$ and parities $P_{J/\psi} = P_\phi = -1$ [2], the CP eigenvalue $n_{J/\psi\phi}$ of the final state depends on the relative angular momentum L, as shown in Equation (2.82).

$$\hat{C}\hat{P}|J/\psi\phi\rangle = n_{J/\psi\phi}|J/\psi\phi\rangle = C_{J/\psi}P_{J/\psi}C_\phi P_\phi \cdot (-1)^L|J/\psi\phi\rangle = (-1)^L|J/\psi\phi\rangle \tag{2.82}$$

The $B_s^0 \to J/\psi\phi$ decay can produce two CP-even states with $L = 0, 2$ and $n_{J/\psi\phi} = 1$ and one CP-odd state with $L = 1$ and $n_{J/\psi\phi} = -1$. The $B_s^0 \to J/\psi\phi$ final state

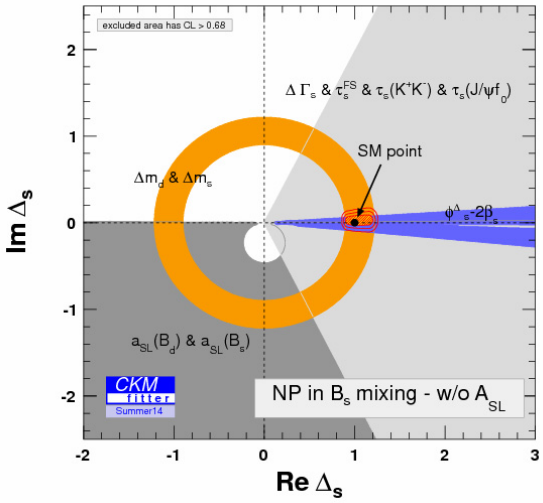


Figure 2.5: Constraints on new physics contributions to B_s^0 mixing in the ($\mathrm{Re}(\Delta_s)$, $\mathrm{Im}(\Delta_s)$) plane. The standard model prediction, $\mathrm{Re}(\Delta_s) = 1$ and $\mathrm{Im}(\Delta_s) = 0$, is compatible with the measurements to 0.3σ. Figure is taken from [35].

selects the heavy or light B_s^0 eigenstate in the initial state, since the CP states correspond to the mass eigenstates of the B_s^0 mesons, as discussed in Section 2.3.4. An angular analysis can be used to identify CP states and hence also the mass eigenstates.

The CP-odd and even components can be disentangled in so called transversity basis [40] that is defined in terms of the momenta $\mathbf{p}_{\mu^+}$, $\mathbf{p}_{\mathrm{K}^+}$, $\mathbf{p}_{\phi(\mathbf{1020})}$ of the positively charged final state muon, kaon, and ϕ meson in the J/ψ rest frame. The coordinate axes of the transversity frame are defined by the unit vectors

$$\begin{aligned}
\hat{\mathbf{x}} &= \frac{\mathbf{p}_{\phi(1020)}}{|\mathbf{p}_{\phi(1020)}|} = \hat{\mathbf{p}}_{\phi(1020)} \\
\hat{\mathbf{y}} &= \frac{\mathbf{p}_{K^+} - \hat{\mathbf{p}}_{\phi(1020)}(\hat{\mathbf{p}}_{\phi(1020)} \cdot \mathbf{p}_{K^+})}{|\mathbf{p}_{K^+} - \hat{\mathbf{p}}_{\phi(1020)}(\hat{\mathbf{p}}_{\phi(1020)} \cdot \mathbf{p}_{K^+})|} \\
\hat{\mathbf{z}} &= \hat{\mathbf{x}} \times \hat{\mathbf{y}}
\end{aligned} \tag{2.83}$$

and the angles are

$$\begin{aligned} \sin\theta\cos\varphi &= \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{x}} \\ \sin\theta\sin\varphi &= \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{y}} \\ \cos\theta &= \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{z}} \\ \cos\psi &= -\hat{\mathbf{p}}'_{K^+} \cdot \hat{\mathbf{p}}'_{J/\psi}, \end{aligned} \tag{2.84}$$

where $\mathbf{p}'_{K^+}$ and $\mathbf{p}'_{J/\psi}$ are the momenta of the positively-charged final-state kaon, and J/ψ meson that are determined in the rest frame of the ϕ meson. The $\hat{x}$-axis is chosen as the flight direction of the ϕ meson, and $\hat{y}$-axis is defined such that $p_y(K^+) > 0$ in the ϕ meson's rest frame. The $\hat{z}$-axis forms a left-handed coordinate system. A schematic picture of the transversity basis with definitions of the transversity angles is shown in Figure 2.6.

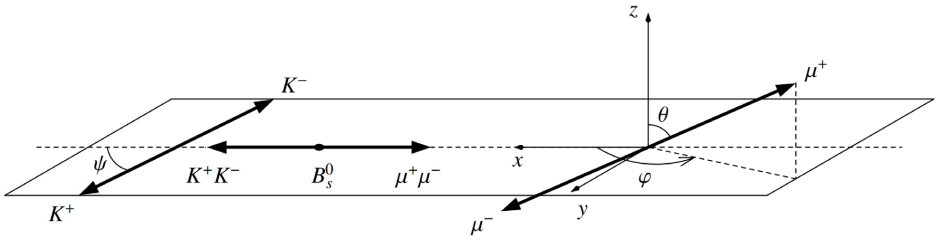


Figure 2.6: Schematic illustration of the transversity basis for $B_s^0 \to J/\psi\phi \to \mu^+\mu^- K^+ K^-$ decay taken from [44].

The differential decay rate for the B_s^0 meson decaying into $J/\psi\phi$ state can be written as

$$\frac{d^4\Gamma\,(B_s^0(t) \to J/\psi\phi)}{d\Theta dt} = f(\Theta, \alpha, t) = \sum_{i=1}^{10} O_i(\alpha, t) \cdot g_i(\Theta), \tag{2.85}$$

where $O_i(\alpha, t)$ are time-dependent functions with a depending set of physics parameters $\alpha_i = \{N_i, a_i, b_i, c_i, d_i\}$. The $g_i(\Theta)$ terms are the angular functions of $\Theta = \{\theta, \psi, \varphi\}$, the decay angles of Equation (2.84) and Figure 2.6.

i	$g_i(\theta,\psi,\varphi)$	N_i	a_i	b_i	c_i	d_i
1	$2\cos^2\psi(1-\sin^2\theta\cos^2\varphi)$	$\lvert A_0(0)\rvert^2$	1	D	C	$-S$
2	$\sin^2\psi(1-\sin^2\theta\sin^2\varphi)$	$\lvert A_\parallel(0)\rvert^2$	1	D	C	$-S$
3	$\sin^2\psi\sin^2\theta$	$\lvert A_\perp(0)\rvert^2$	1	$-D$	C	S
4	$-\sin^2\psi\sin 2\theta\sin\varphi$	$\lvert A_\parallel(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_\parallel)$	$S\cos(\delta_\perp-\delta_\parallel)$	$\sin(\delta_\perp-\delta_\parallel)$	$D\cos(\delta_\perp-\delta_\parallel)$
5	$\frac{1}{\sqrt{2}}\sin 2\psi\sin^2\theta\sin 2\varphi$	$\lvert A_0(0)\rvert\lvert A_\parallel(0)\rvert$	$\cos(\delta_\parallel-\delta_0)$	$D\cos(\delta_\parallel-\delta_0)$	$C\cos(\delta_\parallel-\delta_0)$	$-S\cos(\delta_\parallel-\delta_0)$
6	$\frac{1}{\sqrt{2}}\sin 2\psi\sin 2\theta\sin\varphi$	$\lvert A_0(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_0)$	$S\cos(\delta_\perp-\delta_0)$	$\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\frac{2}{3}(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_S(0)\rvert^2$	1	$-D$	C	S
8	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin^2\theta_T\sin 2\varphi_T$	$\lvert A_S(0)\rvert\lvert A_\parallel(0)\rvert$	$C\cos(\delta_\parallel-\delta_S)$	$S\sin(\delta_\parallel-\delta_S)$	$\cos(\delta_\parallel-\delta_S)$	$D\sin(\delta_\parallel-\delta_S)$
9	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin 2\theta_T\cos\varphi_T$	$\lvert A_S(0)\rvert\lvert A_\perp(0)\rvert$	$\sin(\delta_\perp-\delta_S)$	$-D\sin(\delta_\perp-\delta_S)$	$C\sin(\delta_\perp-\delta_S)$	$S\sin(\delta_\perp-\delta_S)$
10	$\frac{4}{3}\sqrt{3}\cos\psi_T(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_S(0)\rvert\lvert A_0(0)\rvert$	$C\cos(\delta_0-\delta_S)$	$S\sin(\delta_0-\delta_S)$	$\cos(\delta_0-\delta_S)$	$D\sin(\delta_0-\delta_S)$

Table 2.2: Angular and time-dependent terms of the $B_s^0 \to J/\psi\phi$ decay rate model.

The functions $O_i(\alpha, ct)$ are

$$O_i(\alpha,t) = N_i e^{-\Gamma_s t}\left[a_i\cosh(\frac{1}{2}\Delta\Gamma_s t) + b_i\sinh(\frac{1}{2}\Delta\Gamma_s t) + c_i\cos(\Delta m_s t) + d_i\sin(\Delta m_s t)\right], \tag{2.86}$$

where N_i, a_i, b_i, c_i, d_i are given in Table 2.2. Terms 1-6 in Table 2.2 refer to the rate of the P-wave component, i.e. the rate in the $B_s^0 \to J/\psi\phi$ mode. Terms 7-10 describe the non-resonant S-wave component, which refers to a decay rate of the CP-odd $B_s \to J/\psi K^+K^-$ mode. Parameter Γ_s is the average of the heavy and light decay widths defined in Equation (2.49), while $\Delta\Gamma_s$ and Δm_s are the decay width and mass differences of the B_s^0 eigenstates. $|A_\perp(0)|^2$, $|A_0(0)|^2$ and $|A_\parallel(0)|^2$ of Table 2.2 are the squared magnitudes of the perpendicular, longitudinal and parallel amplitudes at $t = 0$ and the parameters $\delta_\perp$, δ_0, $\delta_\parallel$ are their corresponding strong phases. $|A_S(0)|^2$ is the squared magnitude of the S-wave amplitude, and δ_S is its strong phase. It is customary to normalize the amplitudes such that $|A_\perp(0)|^2 + |A_\perp(0)|^2 + |A_0(0)|^2 = 1$. The C, S and D factors are defined as

$$C = \frac{1-|\lambda_{J/\psi\phi}|^2}{1+|\lambda_{J/\psi\phi}|^2}, \qquad S = -\frac{2|\lambda_{J/\psi\phi}|\sin\phi_s}{1+|\lambda_{J/\psi\phi}|^2}, \qquad D = -\frac{2|\lambda_{J/\psi\phi}|\cos\phi_s}{1+|\lambda_{J/\psi\phi}|^2}, \tag{2.87}$$

where ϕ_s is the weak mixing phase. The notations used here were first adopted by

the LHCb collaboration. The model described in Equation (2.85) and Table 2.2 is the decay rate for the B_s^0 meson, whereas the decay rate for $\bar{B}_s^0$ is obtained by switching the signs of the c_i and d_i terms in Table 2.2. Note that we also assume that there is no CP violation in mixing by setting $|\frac{p}{q}| = 1$ in the decay rate of the $\bar{B}_s^0$ mesons. This ratio is measured to be consistent with unity, $|\frac{q}{p}| = 1.0038 \pm 0.0021$ [2] in the B_s^0 system.

2.4.1 Effective lifetime of B_s^0 meson

From now on the discussion concentrates on the P-wave terms of the B_s^0 decay rate as the effective lifetime is defined using only those terms. Hence, the decay rates referred to here only contain the terms 1-6 of Table 2.2. The sum of the decay rates B_s^0 and $\bar{B}_s^0$ mesons to the same state f is

$$\frac{d^4\Gamma}{dtd\Theta} = \frac{d^4\Gamma\,(B_s^0(t) \to f)}{dtd\Theta} + \frac{d^4\Gamma(\bar{B}_s^0(t) \to f)}{dtd\Theta}. \tag{2.88}$$

The c_i and d_i terms in Table 2.2 cancel out since they have opposite signs. The time-dependent functions simplify to

$$O_{i,tot}(\alpha, t) = O_i(\alpha, t) + \bar{O}_i(\alpha, t) = N_i e^{-\Gamma_s t}\left[a_i \cosh(\frac{1}{2}\Delta\Gamma_s t) + b_i \sinh(\frac{1}{2}\Delta\Gamma_s t)\right]. \tag{2.89}$$

In the absence of the direct CP violation, the magnitude of the $\lambda_{J/\psi\phi}$ parameter can be set to unity. This is also experimentally shown by the measurement of the LHCb collaboration, $|\lambda_{J/\psi\phi}| = 0.964 \pm 0.019 \pm 007$ [45]. Furthermore, the mixing phase is both measured and predicted to be small. Using these additional assumptions, the coefficients $C = 0$ and $D \approx -1$ and integrating over the angles, the decay rate in Equation (2.88) can be expressed as

$$\frac{d\Gamma}{dt} = (|A_\parallel(0)|^2 + |A_0(0)|^2)e^{-\Gamma_L t} + |A_\perp(0)|^2 e^{-\Gamma_H t}. \tag{2.90}$$

The magnitudes of the amplitudes at $t = 0$ are from now on denoted by $|A|^2 = |A_\parallel(0)|^2 + |A_0(0)|^2$ and $|A_\perp| = |A_\perp(0)|$. The effective lifetime is defined as the time

expectation value of the total decay rate given by [5],

$$\tau_{eff} = \frac{\int_0^\infty t\frac{d\Gamma}{dt}dt}{\int_0^\infty \frac{d\Gamma}{dt}dt} = \frac{\frac{|A|^2}{\Gamma_L^2} + \frac{|A_\perp|^2}{\Gamma_H^2}}{\frac{|A|^2}{\Gamma_L} + \frac{|A_\perp|^2}{\Gamma_H}}. \tag{2.91}$$

Experimental effects in measurements of the effective lifetime

Experimental biases can affect the value of the measured effective lifetime. First, there is a selection bias, since only B_s^0 candidates with decay times greater than a threshold value a, i.e $t > a$ are considered in order to avoid large resolution and reconstruction effects for short decay times. Second, there is an irreducible background of non-resonant CP-odd $B_s \to J/\psi K^+K^-$ decays among the $B_s^0 \to J/\psi\phi$ candidate events. It is called S-wave bias.

Since the heavy (B_H) and light (B_L) eigenstates have different decay widths, the decay time cut will increase the fraction of the long-lived heavy component B_H in the B_s^0 data. The S-wave bias also increases the fraction of the long-lived CP-odd B_H state in the $B_s^0 \to J/\psi\phi$ sample. Thus, the both effects increase the measured effective lifetime.

The measurement can be corrected for these biases provided their effect on the effective lifetime is quantified. An estimate of the effective lifetime with the decay time cut $t > a$, is obtained by reweighting the two decay rate functions having widths Γ_L and Γ_H. We denote the time integrals of the heavy and light states as C_H and C_L and they can be written as

$$C_H = |A_\perp|^2 \int_0^\infty e^{-\Gamma_H t}dt = \frac{|A_\perp|^2}{\Gamma_H} = |A_\perp|^2\tau_H \tag{2.92}$$

$$C_L = |A|^2 \int_0^\infty e^{-\Gamma_L t}dt = \frac{|A|^2}{\Gamma_L} = |A|^2\tau_L, \tag{2.93}$$

where the lifetimes of the heavy and light components are $\tau_i = 1/\Gamma_i$. In this notation, the effective lifetime in Equation (2.91) becomes

$$\tau_{eff} = \frac{C_H\Gamma_H^{-1} + C_L\Gamma_L^{-1}}{C_L + C_H} = \frac{f_H}{\Gamma_H} + \frac{1 - f_H}{\Gamma_L} = f_H\tau_H + (1 - f_H)\tau_L, \tag{2.94}$$

where the fraction of the heavy component with respect to the total integral is given as $f_H = \frac{C_H}{C_H+C_L} = \frac{|A_\perp|^2\tau_H}{|A_\perp|^2\tau_H+|A_\perp|^2\tau_L}$.

The time integrals of the two decay rate functions must be adjusted when the integration starts from threshold $t = a$ instead of $t = 0$. We then have

$$C'_H = |A_\perp|^2 \int_a^\infty e^{-\Gamma_H t} dt = |A_\perp|^2 \Gamma_H^{-1} e^{-a\cdot\Gamma_H} = |A_\perp|^2 \tau_H e^{-a/\tau_H} \quad (2.95)$$

$$C'_L = |A|^2 \int_a^\infty e^{-\Gamma_L t} dt = |A|^2 \Gamma_L^{-1} e^{-a\cdot\Gamma_L} = |A|^2 \tau_L e^{-a/\tau_L}. \quad (2.96)$$

Replacing $C_{L,H}$ by $C'_{L,H}$ in Equation (2.94) we obtain an estimate for the effective lifetime with the cut $t > a$,

$$\tau_{eff}^{cut} = \frac{C'_H\Gamma_H^{-1} + C'_L\Gamma_L^{-1}}{C'_L + C'_H} = f'_H\tau_H + (1 - f'_H)\tau_L, \quad (2.97)$$

$$f'_H = \frac{C'_H}{C'_H + C'_L} = \frac{|A_\perp|^2\tau_H e^{-a/\tau_H}}{|A_\perp|^2\tau_H e^{-a/\tau_H} + |A|^2\tau_L e^{-a/\tau_L}}.$$

A correction for the S-wave decays is obtained by modifying the fraction of heavy and light components. Denoting the fraction of the S-wave component in the signal events by f_S, the effective lifetime becomes

$$\tau_{eff,S} = (f_P f_H + f_S)\tau_H + f_P(1 - f_H)\tau_L, \quad (2.98)$$

where f_P is the fraction of $B_s^0 \to J/\psi\phi$ decays (P-wave) in the dataset. The fractions of S and P wave decays should add up to one, $f_P + f_S = 1$. To obtain the effective lifetime with a decay time cut, one has to replace f_H with f'_H in Equation (2.98).

2.4.2 Flavour tagging

To reach a maximal sensitivity for the weak mixing phase, it is necessary to distinguish the flavour of the B_s^0 meson at its production. This is seen from the expression of $O_{i,tot}(\alpha, t)$ in Equation (2.89) which depends only on the coefficients a_i and b_i. The weak mixing phase appears only in the form of $\cos\phi_s$ in the total rate, summed over the B_s^0 flavour according to Table 2.2 and expressions (2.87) for the coefficients. On the other hand, $\sin\phi_s$ appears in the coefficients d_i. Sine terms are more

sensitive to the predicted small value of $\phi_s \sim 10^{-2}$ compared to the cosine terms: $\sin\phi_s = \phi_s + O(\phi_s^3)$ and $\cos\phi_s = 1 - \frac{1}{2}\phi_s^2 + O(\phi_s^4)$. It is thus crucial to identify the flavour of the B_s^0 meson at the time of its production.

The flavour tagging is based on the fact that most of the beauty quarks are produced as $b\bar{b}$ pairs. A signal quark hadronizes into a signal B meson, and the other quark forms another B particle. The flavour identification of the signal B meson utilises the particles formed in the hadronization, which are on the same or signal-side (SS) or opposite-side (OS). Several SS and OS flavour tagging techniques have been developed previously by Babar and Belle [41] D0 [46], CDF [47] and LHCb [42] experiments and utilised in the tagging studies presented in this thesis.

Opposite-side tagging

Opposite-side tagging uses the particles created in the hadronization and in decay of the opposite-side b hadron to infer the flavour of the signal meson. The most common OS tagging method is to use electrons and muons from semileptonic decay of the opposite-side b hadron. The charge of the lepton ($l^{\pm}$) and the flavour of the decaying b hadron are correlated: $b \to l^- \bar{\nu}_l c$, $\bar{b} \to l^+ \nu_l \bar{c}$. The charge of the lepton thus determines the flavour of the signal meson. The flavour tagging techniques are schematically shown in Figure 2.7.

There are several complications in the hadronization of the b quarks that affect the tagging performance. Neutral B_d^0 or B_s^0 mesons are created roughly in 50% of the b quark hadronizations [2]. The flavour of the signal meson is challenging to identify in these cases, since the neutral B mesons mix with their antiparticles. Furthermore, the opposite-side b hadron can decay into a final state containing a D meson, $b \to D X$, which then decays semileptonically, $D \to l^+ \nu_l s$. The cascade decays with D mesons have reversed charge-flavour correlation, and hence they lead to wrongly tagged events. About 6% of the tag leptons that originate from b hadrons

are produced in b → c → l chain in the CMS experiment.

Other opposite-side tagging techniques include kaon, vertex- and jet-charge tagging. Opposite-side kaon tagging uses the decay chain b → c → s and searches for a charged kaon created from the strange quark in the cascade decay. As previously discussed, the charge of the kaon correlates with the charge of the signal b hadron. Vertex charge tagging aims at identifying the flavour of the OS b hadron using total charge of the tracks in the decay vertex, and inferring the signal B flavour from the vertex charge. The vertex charge variable is defined as

$$Q_{vtx} = \frac{\sum_i q^i (p_{\mathrm{T}}^i)^k}{\sum_i (p_{\mathrm{T}}^i)^k}, \tag{2.99}$$

where the sum spans over the tracks associated in the OS secondary vertex, q^i is the charge of the i-th track and p_{T}^i is its transverse momentum. Factor k has to be optimised to obtain the best tagging power.

A jet-charge tagger resembles the vertex charge one. The idea of jet charge tagger is to identify the opposite-side b-jet and determine the flavour of the OS b-hadron in a statistical fashion by using charge and p_{T} of the tracks associated with the jet and modifying Equation (2.99) accordingly.

Many factors in a harsh environment of proton-proton collisions and b quark hadronization can dilute the tagging. It is therefore useful to combine as many different tagging strategies as possible to maximize the reliability of the tag decision and tagging efficiency.

Same-side tagging

The same-side tagging techniques utilise particles created in the hadronization of the signal B particle. The most common method is to infer the flavour of the b hadron from the charge of the kaon created in the hadronization of the signal B meson. As the CMS apparatus does not have subdetectors to identify kaons and pions, adapting the same-side kaon tagging will not be as useful as the lepton tagging in

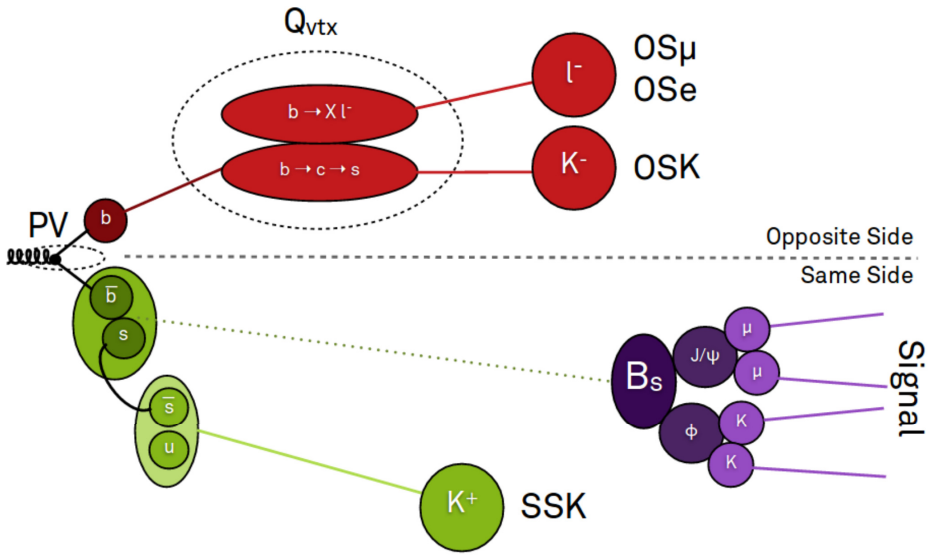


Figure 2.7: Schematic view of the opposite and same-side tagging methods. PV represents the production vertex of the $b\bar{b}$ pair. The opposite-side vertex charge Q_{vtx}, kaon (OSK) and lepton taggers (OSe, OSμ) are shown with same-side kaon tagger (SSK). $B_s^0 \to J/\psi\phi$ acts as the signal decay. Figure is taken from [48].

the CMS experiment.

Characterisation of the tagging performance

Three measures characterise the performance of a tagging algorithm: the mistag fraction, tagging efficiency and tagging power. Mistag fraction is the probability of misidentifying the flavour of the meson and is defined as

$$\omega = \frac{N_W}{N_W + N_R}, \tag{2.100}$$

where N_W and N_R are numbers of wrongly and correctly tagged B mesons. The tagging efficiency ϵ is the probability of finding a tagger for determining the flavour of the meson,

$$\epsilon_{tag} = \frac{N_W + N_R}{N_W + N_R + N_U}, \tag{2.101}$$

where N_U refers to the number of untagged B meson events. The tagging power is a figure of merit of the tagging algorithm given by a product of tagging efficiency and mistag fraction

$$P_{tag} = \epsilon_{tag} D^2 = \epsilon_{tag}(1-2\omega)^2, \tag{2.102}$$

where D stands for dilution, $D = 1-2\omega$. The effective statistical size of a tagged B meson sample with N events is equivalent to a perfectly tagged sample with $N\epsilon D^2$ events.

The tagging performance can also be determined in mutually exclusive categories such as in bins of lepton p_T. In this case the tagging performance is

$$P_{tag} = \sum_{i=1}^{N_{bins}} \epsilon_{tag}^i (1-2\omega^i)^2, \tag{2.103}$$

where ϵ_{tag}^i and ω^i are the efficiency and wrong tag fraction in the i-th bin. In this case, the total efficiency is defined as

$$\epsilon_{tag} = \sum_{i=1}^{N_{bins}} \epsilon_{tag}^i. \tag{2.104}$$

The average mistag fraction ω is obtained from

$$\omega = \frac{1}{2}\left(1 - \sqrt{\frac{P_{tag}}{\epsilon_{tag}}}\right), \tag{2.105}$$

where P_{tag} is the tagging power from Equation (2.103) and ϵ_{tag} is the total efficiency defined in Equation (2.104).

The mistag fraction can also be parametrised using the binned wrong tag fraction distribution. The tagging power is then

$$P_{tag} = \frac{1}{N_{tot}} \sum_{i=1}^{N_{tagged}} (1-2\omega^i)^2, \tag{2.106}$$

where N_{tot} is the total number of signal events and the sum is over the tagged signal events N_{tagged}. The average mistag fraction associated with the parametrised tagging performance can be evaluated from Equation (2.105), where the total efficiency is obtained from Equation (2.101).

Decay rate with mistag information

The rate models, including the event-by-event mistag information, are the sums of the differential decay rates of B_s^0 and $\bar{B}_s^0$ mesons weighted by the mistag fraction. The models for B_s^0 and $\bar{B}_s^0$ become

$$R\left[B_s^0 \to f\right] = (1-\omega^i)\frac{d\Gamma\left(B_s^0(t)\to f\right)}{N_f dt} + \omega^i \frac{d\Gamma(\bar{B}_s^0(t)\to f)}{N_f dt} \tag{2.107}$$

$$R\left[\bar{B}_s^0 \to f\right] = \omega^i \frac{d\Gamma\left(B_s^0(t)\to f\right)}{N_f dt} + (1-\omega^i)\frac{d\Gamma(\bar{B}_s^0(t)\to f)}{N_f dt}. \tag{2.108}$$

By assuming that $|\frac{p}{q}| = 1$ and $|\lambda_{J/\psi\phi}| = 1$ and adding up the differential rates defined in Equations (2.63) and (2.64), the decay rate models simplify to

$$\begin{aligned} R\left[B_s^0 \to f\right] = |A_f|^2 e^{-\Gamma_s t}\{&\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - n_f \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\cos\phi_s \\ &+ n_f(1-2\omega^i)\sin\left(\Delta m_s t\right)\sin\phi_s\} \end{aligned} \tag{2.109}$$

$$\begin{aligned} R\left[\bar{B}_s^0 \to f\right] = |A_f|^2 e^{-\Gamma_s t}\{&\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - n_f \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\cos\phi_s \\ &- n_f(1-2\omega^i)\sin\left(\Delta m_s t\right)\sin\phi_s\}. \end{aligned} \tag{2.110}$$

Note that the real and imaginary parts of the CP violation parameter λ_f in Equations (2.63) and (2.64) are denoted here as $\mathrm{Re}(\lambda_{J/\psi\phi}) = n_f \cos\phi_s$ and $\mathrm{Im}(\lambda_{J/\psi\phi}) = n_f \sin\phi_s$, where n_f is the CP eigenvalue of the $J/\psi\phi$ state. Moreover, the rate models in Equations (2.109) - (2.110) can be expressed using the tag decision q_i

$$\begin{aligned} R\left[B_s^0\ (\bar{B}_s^0) \to f\right] = |A_f|^2 e^{-\Gamma_s t}\{&\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - n_f \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\cos\phi_s \\ &+ q_i n_f(1-2\omega^i)\sin\left(\Delta m_s t\right)\sin\phi_s\}, \end{aligned} \tag{2.111}$$

where $q_i = +1$ represents the meson tagged as B_s^0, $q_i = -1$ indicates the meson tagged as $\bar{B}_s^0$, and for the untagged event $q_i = 0$. The rate presented in Equation (2.111) together with the angular distributions is used in the experimental likelihood model to measure the weak mixing phase ϕ_s.

2.4.3 Mixing asymmetry

The mistag fraction and the performance of a tagging algorithm can be measured using the B_d^0 data and the mixing asymmetry, defined by

$$A_{mix}(t) = \frac{P_{unmix}(t) - P_{mix}(t)}{P_{unmix}(t) + P_{mix}(t)} \tag{2.112}$$

where $P_{mix}(t) = P^{B_d^0}_{mix}(t) + P^{\bar{B}_d^0}_{mix}(t)$ and $P_{unmix}(t) = P^{B_d^0}_{unmix}(t) + P^{\bar{B}_d^0}_{unmix}(t)$ are the sums of probabilities that a neutral B meson does or does not oscillate into its antiparticle. The mixing probabilities are given in Equations (2.59) - (2.60).

The world-average value of the decay width difference $\Delta\Gamma$ for B_d^0 system is consistent with zero, as shown in Table 2.1, and the ratio of mixing parameters $|\frac{q}{p}|$ is measured to be consistent with unity, $|\frac{q}{p}| = 1.0007 \pm 0.0009$ [2]. Using these experimental results, the sums of mixing probabilities can be approximated as

$$P_{mix}(t) \quad = P^{B_d^0}_{mix}(t) + P^{\bar{B}_d^0}_{mix}(t) \approx e^{-\Gamma t}\,(1 - \cos(\Delta m t)\ ) \tag{2.113}$$

$$P_{unmix}(t) \quad = P^{B_d^0}_{unmix}(t) + P^{\bar{B}_d^0}_{unmix}(t) \approx e^{-\Gamma t}\,(1 + \cos(\Delta m t)\ )\,. \tag{2.114}$$

The mixing asymmetry in Equation (2.112) then simplifies to

$$A_{mix} \approx D\cos(\Delta m t), \tag{2.115}$$

where amplitude D is the dilution $D = 1 - 2\omega$. Theoretically, the dilution equals to one, but experimentally it is always smaller than unity due to the mistagging of the initial or final B_d^0 meson flavour.

3. Proton-proton collisions

This chapter reviews the basic physics concepts related to the proton-proton collisions such as luminosity and rapidity. Furthermore, the Large Hadron Collider, its acceleration chain and physics experiments are briefly described.

3.1 Physics concepts

3.1.1 Luminosity

Luminosity is a measure that describes the particle flux traversing through a square centimeter area per second. A good approximation for instantaneous luminosity $\mathcal{L}_{inst}$ for two beams with a Gaussian beam profile is [49]

$$\mathcal{L}_{inst} = f\frac{N_1 N_2 N_b^2}{4\pi\sigma_x\sigma_y}, \tag{3.1}$$

where N_b is the number of bunches in one beam, $N_{1,2}$ are the numbers of particles in the bunches, f is the rotation frequency and $\sigma_{x,y}$ are the root-mean-square widths of the bunches in x and y directions.

By integrating the instantaneous luminosity over time when the collisions are ongoing, one obtains integrated luminosity $\mathcal{L}_{int}$. The integrated luminosity is often expressed in units of inverse barns: 1 barn = 10^{-28} cm^{-2}. By multiplying it with a cross section of a process $\sigma_{process}$ expressed in units of barns, one obtains the expected number of events N_{events} from the process

$$N_{events} = \mathcal{L}_{int}\sigma_{process}. \tag{3.2}$$

The cross section is a measure of an occurrence probability of the process.

The number of expected events N_{events} can be increased by increasing the instantaneous luminosity. From Equation (3.1), it is easy to see that instantaneous luminosity can be also increased by increasing the numbers of the colliding particles (N_1, N_2) in beams, the number of bunches (N_b) or by squeezing the bunches to a smaller width (σ_x, σ_y). For some processes, e.g. for the production cross section of $b\bar{b}$ pair, the increase in beam energy will also enlarge the cross section. However, the increase in energy will not always enlarge the cross section, and this is only true for some physics processes.

3.1.2 Rapidity and pseudorapidity

The interacting particles at the TeV energies are not protons but their constituting quarks and gluons. The partons carry a fraction of the proton's momentum along the z-axis. The system of two interacting partons has often a sizeable net momentum in the experimental reference frame. Therefore the particles created in the interaction are relativistic and boosted along the beam direction. Thus, observables that are invariant under Lorentz transformation are considered useful. One such measure is rapidity (see e.g. [50]),

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.3}$$

where E is the energy of the particle and p_z is the projection of the momentum along the beam axis. The rapidity is also an additive quantity unlike particle velocities in the relativistic regime. In the ultra-relativistic limit, where the momentum of the particle is much bigger than its mass, $p >> m$, rapidity can be approximated to pseudorapidity. Pseudorapidity (see e.g. [50]) is

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right), \tag{3.4}$$

where is θ the angle between the total three momentum p and the momentum along the z-axis, $\cos\theta = \frac{p_z}{p}$.

3.2 The Large Hadron Collider

The Large Hadron Collider [51] at CERN is currently the biggest particle accelerator in the world. The LHC is located on the border of France and Switzerland, near the city of Geneva. It is a double-synchrotron ring with a 26.7 km circumference located underground at a depth of 100 m. The LHC accelerates two beams of protons and its design collision energy is 14 TeV. During 2010 - 2017 the collision energies of 7,8 and 13 TeV have been used. The design luminosity of the LHC is 10^{34} cm^{-2}s^{-1}.

3.2.1 Acceleration chain

The LHC uses ionized hydrogen as a source of protons. The protons are pre-accelerated in a series of smaller accelerators before they are injected into the LHC ring. The beam first enters Linac2 where it gains energy of 50 MeV. The protons obtain more energy in the PS Booster (PSB, 1.4 GeV), in the Proton Synchrotron (PS, 25 GeV) and in the Super Proton Synchrotron (SPS, 450 GeV) after which they are injected into the LHC. A schematic image of the acceleration chain is shown in Figure 3.1.

At SPS, the protons are divided into two counter-rotating beams that contain in total 2808 bunches of protons. Each bunch comprises approximately 10^{11} protons and bunch spacing is 25 ns or 50 ns, depending on the year when the collisions are recorded.

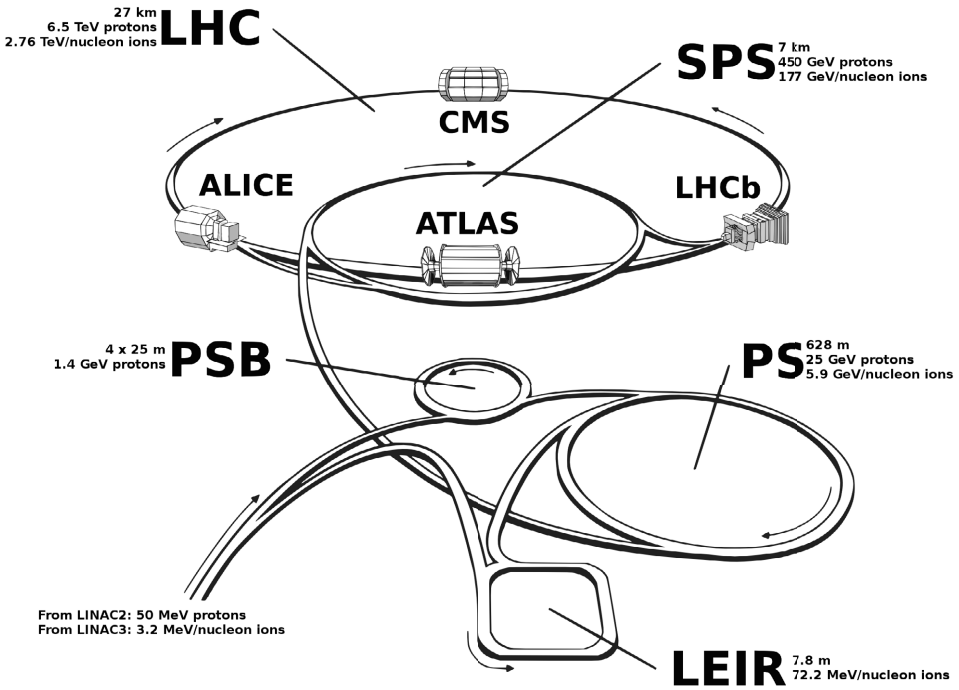


Figure 3.1: The LHC accelerator chain. The acceleration begins with Linac2. The protons then obtain more energy in the PS Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), after which they are injected into the LHC. Energies of protons at each acceleration step are shown on the left side of accelerator names. Figure is taken from [52].

3.2.2 Magnets

The LHC has 1232 dipole magnets that are used to circulate the protons around the LHC ring. The magnets comprise superconducting niobium-titanium (NbTi) coils that are cooled to the temperature of 1.9 K using liquid helium. The magnetic field strength of the dipoles can reach 8.3 T. In addition to the dipole magnets, the LHC employs 392 quadrupole magnets for beam focusing. Other magnets with larger numbers of poles are also deployed to correct for the scattering of protons from the beam orbit.

3.2.3 The LHC experiments

There are four big experiments at the LHC: ALICE, ATLAS, CMS and LHCb. The CMS and ATLAS experiments are designed to study a wide range of physics phenomena from low-energy B physics to high-energy regime of Higgs-physics, dark matter, and supersymmetry. The LHCb apparatus is devoted to the studies of b quarks, b hadrons and CP violation. The ALICE experiment is designed for the detection of quark gluon plasma, a dense matter where quarks are not bound into the hadrons. The ATLAS, the CMS and the LHCb experiments primarily analyse the data recorded from proton-proton interactions, whereas the main interest of the ALICE collaboration is collisions of lead ions.

The four large experiments are complemented by three smaller experiments, TOTEM, LHCf [53] and MoEDAL [54]. The TOTEM experiment is designed to measure the total cross section of proton-proton collisions in the forward region. The LHCf experiment studies production of neutral particles in the forward region to improve modelling of cosmic rays interactions in the atmosphere. The MoEDAL collaboration searches for magnetic monopoles created in proton-proton collisions.

4. The Compact Muon Solenoid

The Compact Muon Solenoid [1] is a cylindrical detector system that records the signals of particles originating from the proton-proton collisions. The overall length of the detector is 28.7 m, and it has a diameter of 15.0 m. The CMS apparatus weighs 14000 t. The innermost parts of the experimental setup are the silicon pixel and silicon strip tracker. These detectors are used to record the trajectories of charged particles that traverse through the pixel and strip layers. The tracker is followed by the electromagnetic calorimeter (ECAL), which is constructed from lead-tungstate crystals. ECAL is designed for measuring the energy of photons and electrons. ECAL is surrounded by the hadronic scintillator calorimeter (HCAL), which measures the energies of charged and neutral hadrons. The heart of the CMS detector is a superconducting solenoid that produces the magnetic field of 3.8 T. The magnet is positioned between the hadronic calorimeter (HCAL) and the return yoke. The muon chambers embedded in the return yoke are used to reconstruct the muon tracks and to measure their transverse momentum. The forward region near the beam pipe is covered by the forward subdetector of the HCAL. A picture of the CMS detector appears in Figure 4.1.

The CMS experiment uses a right-handed coordinate system whose origin is set to the interaction point. The x-axis points towards the center of the LHC ring and the y-axis points upwards at the sky. The z-axis is set counter-clockwise along the direction of the proton beam. A cylindrical coordinate system is also used. The

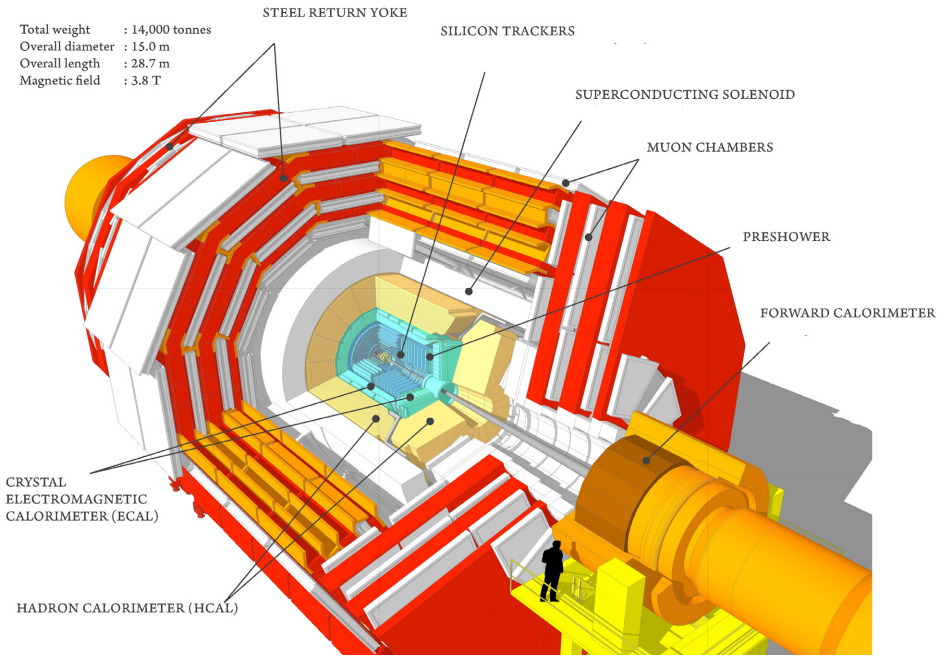


Figure 4.1: A schematic view of the CMS detector.

azimuthal angle ϕ is measured in the x - y plane while the polar angle θ is defined from the positive z-axis.

4.1 Tracking system

The tracking system [1, 55, 56] of the CMS apparatus consists of two subdetectors: the silicon pixel tracker and the silicon strip tracker. The tracking system detects spatial coordinates of charged particles traversing through the tracker layers. The positions of the primary interaction vertices of the colliding protons can be identified using the tracks reconstructed in the tracker. Moreover, the secondary vertices of the short-living particles decaying in the beam pipe can also be extrapolated from the particle trajectories. A schematic view of the tracker is shown in Figure 4.2.

4.1.1 Pixel tracker

The silicon pixel tracker consists of three cylindrical layers of pixel detector modules in the barrel and two endcap disks on both sides of the barrel. The three barrel layers having length of 53 cm are located at the radial distances of 4.3 cm, 7.3 cm and 10.2 cm from the beam axis. The endcap disks are positioned at the radii of 4.8 cm and 14.4 cm from the beam axis and at the longitudinal distance of 35.5 cm and 48.5 cm from the interaction point. The three barrel layers cover the pseudorapidity region of $|\eta| < 2.2$.

The total surface area of the pixel tracker is 1 m^2 and it comprises 66 million pixel sensors that have a size of 100×150 μm^2. The pixel sensors are pn junctions working at an n-on-n concept: High-dose n^+-type electrodes are embedded in a high-resistance n-type substrate and the p-type implant is attached on the backside of the sensor. The charge carriers of the sensors are electrons that are exposed to the Lorentz force due to the high magnetic field of the CMS apparatus. The Lorentz force changes the drift direction of the electrons which are collected by more than one p-type implant. To enhance the charge sharing of the neighbouring sensors at the endcap disks, the forward detectors are tilted at 20° forming a turbine-like geometry. The charge sharing improves the spatial resolution of the pixels down to 10 μm in the $r - \phi$ direction and 20 μm in the z-direction.

4.1.2 Strip tracker

The silicon microstrip tracker comprises four individual subdetectors. The barrel region is divided into the tracker inner barrel (TIB) and the tracker outer barrel (TOB) whilst the endcap plates are grouped into tracker inner disks (TIDs) and tracker endcaps (TECs).

The TIB detector contains four layers of silicon microstrip modules that have strip spacing of 80 μm in the two innermost layers and strip spacing of 120 μm in

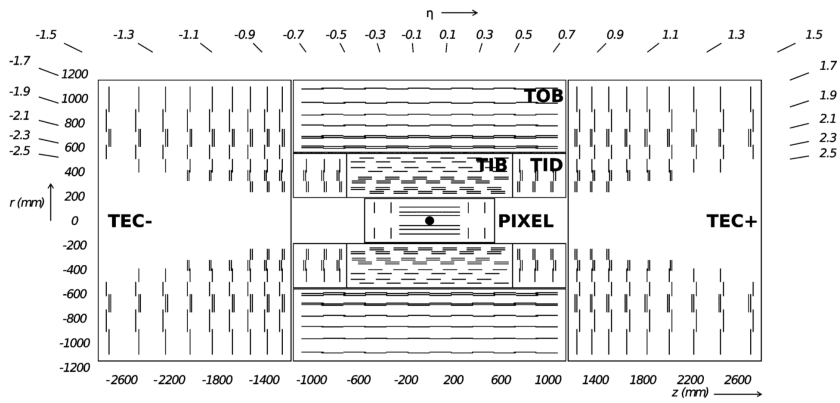


Figure 4.2: A schematic view of the CMS tracker in $r - z$ plane. Figure is taken from [56].

the two outermost layers. The TIB layers are located at radial distances of 25.5 cm, 33.9 cm, 41.8 cm and 49.8 cm from the beam axis and extend from -70 cm to 70 cm around the interaction point. The position resolution of the inner barrel strip tracker is 23-34 μm in the $r - \phi$ plane and 23 μm in the z direction. There is in total of six inner disks in the strip tracker, three at both ends of the TIB. The disks are placed between $\pm$80 cm and $\pm$90 cm in the beam axis. The TID has the strip pitch of 100 -141 μm depending on the disk.

The TOB consists of six cylindrical layers of detector modules. The layers are positioned between the inner and the outer radii of 55.5 cm and 116.0 cm, respectively. The strip pitch in the TOB is between 120-180 μm depending on the layer. The position resolution of the outer barrel layers is 35-52 μm in the $r - var\phi$ plane and 52 μm in z-axis. The nine disks of tracker endcaps range from 124 cm to 280 cm in the beam axis. The endcap disks have a strip pitch of 97-184 μm.

The strip tracker is composed of 15148 detector modules that are made of one or two sensors. The sensor comprises p^+ type strips attached to a phosphorus-doped n-type substrate. An active area of the modules varies from 6245 mm^2 in the tracker endcaps to 17202 mm^2 in the outer barrel.

4.2 Calorimeters

The electromagnetic calorimeter [55, 57] measures the total energy of electrons and photons and is used to in the jet reconstruction. ECAL extends from radii of 116 cm to 177 cm from the beam axis. The calorimeter consists of 61200 lead tungstate ($PbWO_4$) crystals in the barrel and 7324 crystals at each of the endcaps. The cross-section of the crystals varies between 2.2 x 2.2 cm^2 and 2.6 x 2.6 cm^2 depending on their placement in the calorimeter. The crystals have a length of 22-23 cm. The pseudorapidity coverage of ECAL is $|\eta| < 3$. The electromagnetic calorimeter is presented in Figure 4.3.

When an electron or a photon enters ECAL, it creates a cascade of secondary photons, electrons, and positrons. The shower traverses through the detector volume creating fluorescent light that is detected by avalanche photodiodes in the barrel and vacuum phototriodes at the endcaps.

Hadronic calorimeter [55] detects the hadrons that live long enough to reach the calorimeter and is used to reconstruct jets. HCAL is divided into four subdetectors: hadronic forward (HF), hadronic barrel (HB), hadronic outer (HO), and hadronic endcap (HE) calorimeters. The barrel calorimeter extends from 1.77 m to 2.95 m radius from the beam axis while HO is placed in the radial distance of 3.82 m to 4.07 m from the beam axis. The forward calorimeter is located at 11.2 m from the collision point. The coverage of the barrel calorimeter and the endcaps is up to $|\eta| < 3$.

The HCAL barrel and endcaps are sampling calorimeters that comprise alternating layers of absorbers and three types of plastic scintillators ordered in trapezoidal towers. The brass and steel plates create a cascade of secondary particles. The shower produces fluorescent light in the scintillator layers. The light is guided to wavelength shifting fibres and detected by hybrid photodiodes.

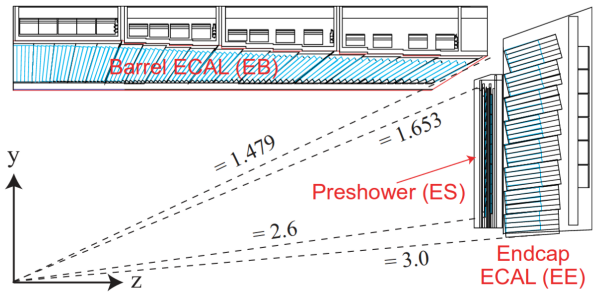


Figure 4.3: An illustration of the CMS electromagnetic calorimeter. Figure is taken from [1].

4.3 Magnet

The experiments in high-energy physics use a magnetic field to enable a momentum measurement of a charged particle. When a charged particle passes through a magnetic field, its trajectory curves due to the Lorentz force. The momentum can be determined from the sagitta of the curved track. The trajectories with opposite charges curve in opposite directions. This allows identification of the charge of the particle.

The superconducting solenoid [1] of the CMS experiment has been designed for 4 T magnetic flux density but it is being operated at 3.8 T field. The magnet has a diameter of 6 m and a length of 12.5 m. The solenoid and the return yoke weigh 10000 t. The purpose of the return yoke is to increase the homogeneity of the field in the tracker and reduce the stray field. The return yoke is constructed from five wheels and two endcaps that consist of three disks each. The solenoid operates at 4.5 K temperature and has a nominal current of 19.14 kA.

4.4 Muon detectors

The muon spectrometer is designed to identify muons and measure their momenta. Muon chambers are also an important part of the CMS triggering system. The cham-

bers are composed of three types of gaseous detectors: drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The muon spectrometer has coverage of up to $\eta = 2.4$ and is depicted in Figure 4.4.

4.4.1 Drift tubes

The drift tubes [58] occupy the barrel of the muon system and cover the pseudo-rapidity region $|\eta| < 1.2$. The system extends from a radial distance of 4.02 m to 7.38 m from the beam axis. The DTs are organised into four stations (MB1-MB4) that are embedded in the flux return yoke of the solenoid. The stations are divided into twelve sectors in the transverse plane. Each station has one DT chamber per section except the uppermost and lowermost sectors 4 and 10, which consist of two DT chambers. There are twelve layers of drift cells inside a DT chamber.

The rectangular drift cells have a cross section of 1.3 x 4.2 cm^2. The heart of the cell is an anode wire. Two aluminium electrodes below and above the anode wire and a cathode strip shape the electric field lines in the drift cell. The drift cell operates with a gas mixture of 85% Argon and 15% carbon dioxide. Each drift cell reconstruct a muon coordinate by measuring the drift time of the ionization electrons to the anode wire and then converting the drift time into a distance from the wire. Position resolution of the drift tube is 100 μm in $r - \phi$ plane and 100-150 μm in $r - z$ plane.

4.4.2 Resistive plate chambers

The resistive plate chambers [59] are placed in the barrel and at the endcaps. RPCs are used for triggering and they also deliver fast timing information for CSCs and DTs about the muon entering and leaving the muon system. This information can be used to convert to drift time of electrons and ions measured in the DTs or CSCs into position coordinates.

The RPCs consist of two resistive parallel plates separated by a 2 mm gap. One of the plates serves as an anode and the other plate as a cathode. The volume between the plates is filled with a mixture that consists of 95.2% Freon, 4.5% isobutane and 0.3% sulphur hexafluoride. The spatial resolution of the RPCs is an order of a centimeter that corresponds to a size of a read out strip.

4.4.3 Cathode strip chambers

The cathode strip chambers [60] are installed in the endcaps of the CMS apparatus and grouped into four disks labelled as ME1-ME4. A basic units of the CSCs are multiwire proportional counters that are organised in six instrumenting layers. Each layer has an anode wire between the two planar copper cathodes. One of the cathodes is segmented into seven strips to provide a position measurement and another cathode is continuous. The CSCs are filled with a gas mixture of 40% Argon, 50% CO_2 and 10% CF_4. The CSCs can measure both orthogonal and longitudinal crossing positions of the muons. Spatial resolution of the CSCs is 75 or 150 μm, depending on the station.

4.5 Triggering system

The triggering system is a key component of particle physics experiments as the data given by the detectors has to be processed and stored quickly. With a bunch crossing interval of 25 ns, collisions take place at a rate of 40 MHz. Since the collisions occur at a high rate and the size of the data per event is about a megabyte, it is impossible to collect and save the data from every collision event delivered by the LHC.

An efficient triggering system is developed to select interesting collision events. The triggering is divided into two stages: the first-level (L1) trigger [61] and the high-level trigger (HLT) [62]. The L1 trigger consists of hardware processors that use the

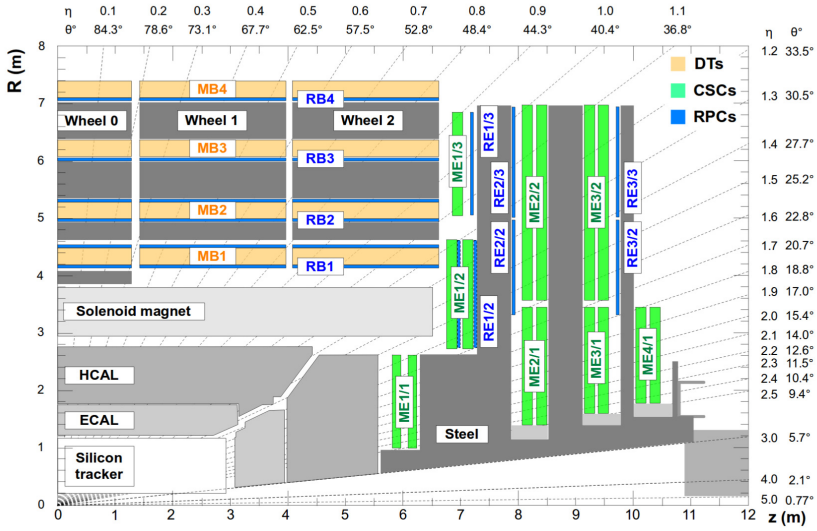


Figure 4.4: A schematic view of the CMS muon system in $r - z$ plane. Figure is taken from [55].

information coming from the calorimeters and the muon chambers to identify and store the data from the most interesting events. The time spent to select or reject the event is 3.2 μs. The event selected by the L1 trigger is directed to the HLT processor farm. The HLT uses more sophisticated event reconstruction methods to deliver the final decision as to whether to store the event. The L1 trigger reduces the event rate from 40 MHz to 100 kHz and the HLT further decreases the rate from 100 kHz to 100-300 Hz. The stored datasets are classified according to physics criteria defined by the L1 and high-level triggers.

4.5.1 L1 trigger

The L1 trigger identifies muons, jets, electrons/photons and missing transverse energy based on the regional information of calorimeters and muon detectors. It has three subsystems: the L1 calorimeter trigger, L1 muon trigger, and L1 global trigger. The L1 calorimeter trigger uses the information coming from the ECAL and HCAL towers, whilst the L1 muon trigger utilises the information from DTs, CSCs, and

RPCs. The L1 global trigger combines the output of the calorimeter and the muon triggers and makes the final trigger decision. A schematic view of the L1 trigger system is illustrated in Figure 4.5.

The L1 calorimeter trigger consists of a regional calorimeter trigger (RCT) and a global calorimeter trigger (GCT). The RCT determines jets, electron/photon candidates and their transverse energy in the different regions of the calorimeters. Each region has a size of 4 x 4 trigger towers for ECAL and HCAL and one tower for HF.

The RCT sends the sums of transverse energy and the jet and electron/photon candidates to the GCT that classifies central, forward and tau jets. It calculates the missing transverse energy and the scalar transverse energy sum (H_T) of all jets above a certain threshold. The GCT sorts the electron/photon candidates and each types of jets according to the energy and the quality of the energy measurement and sends the objects to the global trigger.

The L1 muon trigger is composed of regional muon triggers and the global muon trigger (GMT). Regional triggers for CSC and DT subsystems are further divided into trigger primitives that identify the track segments and higher level track finders that form the complete muon tracks. The RPC subdetector has a single pattern comparator system that compares signals from all four muon stations to predefined hit patterns in order to find muon candidates. Both the CSC and DT track finders deliver up to four best objects to the global muon trigger. The RPC pattern comparator sends a maximum of eight best muons to the GMT, four in the barrel, and four in the endcaps. The information given to the GMT includes track transverse momentum, η and ϕ, muon charge and the quality code assigned to the track.

The GMT combines candidates identified by more than one subdetector and assigns each track a global quality code based on the number of subdetectors involved,

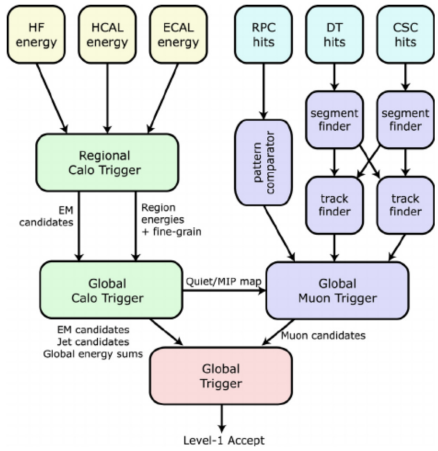


Figure 4.5: A schematic illustration of the L1 trigger decision chain. Figure is taken from [55].

and the quality score previously determined by the regional trigger. A maximum of four highest-quality muon candidates are forwarded to the global trigger.

The global trigger first synchronises the muon and calorimeter input data and transmits them to a global trigger logic (GTL) board. The board contains a trigger menu of up to 128 algorithms, which transform the logical combinations L1 trigger objects such as muons, jets, and missing energy into decision bits. These bits contribute to a final decision as to whether the event is given to the HLT or rejected.

4.5.2 HLT trigger

The full information of the events accepted by the L1 trigger is directed to the software based high-level trigger. The time for the decision-making at the HLT level is at maximum of 300 ms. This sets a constraint on the complexity of the event reconstruction. In order to reject uninteresting events as quickly as possible, the HLT has three virtual layers that add more information on the event reconstruction. The level-2 layer uses the information from calorimeters and the muon detectors,

level-2.5 includes partial information from the tracker (e.g. pixel hits) and level-3 reconstructs the full particle trajectories in the tracker. The event has to pass each of the layers in order to be accepted by the HLT.

The track reconstruction is the most time-consuming step of the high-level trigger. Thus, it is performed only in the regions of interest indicated by the L1 trigger and using a limited set of hits. The event reconstruction and a selection of various physics signatures, such as J/ψ mesons displaced from the beam line, is organised in trigger paths. The paths contain the three virtual layers of the event reconstruction followed by a set of selection criteria after each reconstruction step.

5. Event reconstruction

The reconstruction of the particles produced in the collisions is done iteratively starting from signals in single subdetector sensors and combining them to create more complex objects such as trajectories of charged particles. In this chapter, an overview of the particle reconstruction and categorization is given for those physics objects that are relevant to the analyses presented in the thesis.

5.1 Tracks

Charged particles such as muons and charged hadrons interact with the sensors in pixel and strip tracker and deposit a small amount of their energy inside them. These deposits are called tracker hits. The trajectory of a charged particle is a helix assuming a uniform magnetic field in the tracker and neglecting the small energy losses of the particles.

Five parameters describe a particle track in the context of the CMS software [63]: d_0, z_0, ϕ, $\cot\theta$, and transverse momentum p_{T}. The track parameters are determined at the point of closest approach with respect to the beam axis (x_0, y_0, z_0), which is named as an impact point. The d_0 and z_0 define the impact point position in the radial and z directions, $d_0 = -y_0 \cos\phi + x_0 \sin\phi$. The azimuthal and polar angels of the track momentum vector are given by ϕ and θ, respectively. The track reconstruction aims to find the hits belonging to a single track and then fits the best possible values for the track parameters.

Track reconstruction of the CMS experiment uses the combinatorial track finder (CTF) algorithm [63] that is based on the Kalman filter [64] technique. The track reconstruction proceeds in five steps: i) fast and coarse reconstruction of pixel vertices ii) generation of the track seeds iii) track finding iv) track fitting and v) track selection and categorisation. The reconstruction steps i-iv are repeated in six track iteration rounds. Each iteration reconstructs different types of tracks. The hits associated with the reconstructed tracks are removed from the hit collection after each iteration.

The tracking starts with a coarse reconstruction of primary vertices using track stubs that are formed of three pixel hits. The reconstruction of the pixel vertices is followed by the generation of the track seeds. The track seeds are formed by the combinations of three pixel hits in different pixel layers or two pixel hits having a compatibility with one of the pixel vertices. The seeds are then extended towards the outermost layers of the tracking detector using the Kalman filter algorithm. The algorithm searches for hits compatible with the original track seed in a cone whose size is proportional to the precision of the track parameters. A new track is formed each time when multiple hits consistent with the search window are found. Thus, each track seed can have several tracks associated with it. Track finding is terminated when the last tracker layer is reached or when no hits are available for the track reconstruction.

The tracks associated with a track seed are fitted using a χ^2 technique. For each track seed, a maximum of five tracks having the largest fit probabilities are retained. More hits that belong to the same trajectory are searched for by propagating the track from the outermost hit inwards. This is called track smoothing. Finally, a helix is fitted to the track hits in order to obtain the final estimates of the track parameters.

After finishing the track iteration, a track cleaner algorithm goes through all

pairs of track candidates and calculates the fraction of shared hits. If the track pair shares more than 19% of hits, the track with fewer hits is removed. The tracks are then grouped into different categories like high-purity tracks or loose tracks based on quality measures such as number of hits associated with the track, number of tracker layers used in the track reconstruction, and χ^2/dof of the track fit.

The first track iteration reconstructs prompt tracks that have $p_{\mathrm{T}} > 0.8$ GeV and three hits in the pixel detector. The second iteration concentrates on the prompt tracks having two pixel hits and $p_{\mathrm{T}} > 0.6$ GeV. The third iteration round searches for low-p_{T} tracks, and the remaining three searches reconstruct tracks that originate outside of the luminous region (beam spot) of the beam axis.

The last three iterations are particularly interesting for the B physics analyses as the tracks that originate from the decays of b hadrons are often displaced from the beam axis. This happens because of the long average lifetime of b hadrons. Tracking efficiency for different track iterations is shown in Figure 5.1.

The relative transverse momentum resolutions for muons and pions are shown in Figure 5.2. For a typical muon with $p_{\mathrm{T}} = 10$ GeV, the relative p_{T} resolution is about 1%. For a pion of $p_{\mathrm{T}} = 1$ GeV in the pseudorapidity region $|\eta| < 1$, the relative resolution is within 0.8-2%. In the measurements using the b hadron decay time, the p_{T} resolution of the b hadron decay products is one of the ingredients affecting the proper decay time resolution.

5.1.1 Track categorisation

The CMS experiment has developed different categories of tracks that are tailored to a variety of physics analyses. There are loose, tight, and high-purity track collections [65]. Here, only the high-purity track selection is described in more detail, since it is used in the $\mathrm{B}_{\mathrm{s}}^{0}$ effective lifetime measurement.

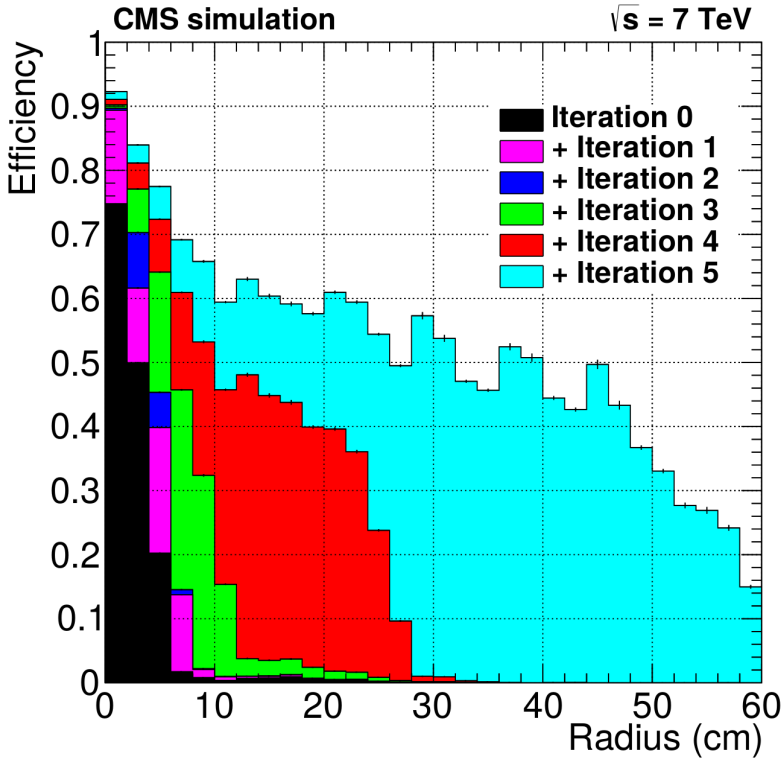


Figure 5.1: Tracking efficiency as a function of transverse distance from the beam axis to the production point of each particle for the different track reconstruction iterations. The tracks are required to pass the high-purity quality requirement and have $p_{\mathrm{T}} > 0.9$ GeV and $|\eta| < 2.5$. The efficiency is evaluated using simulated $\mathrm{t\bar{t}}$ events. Figure is taken from [65].

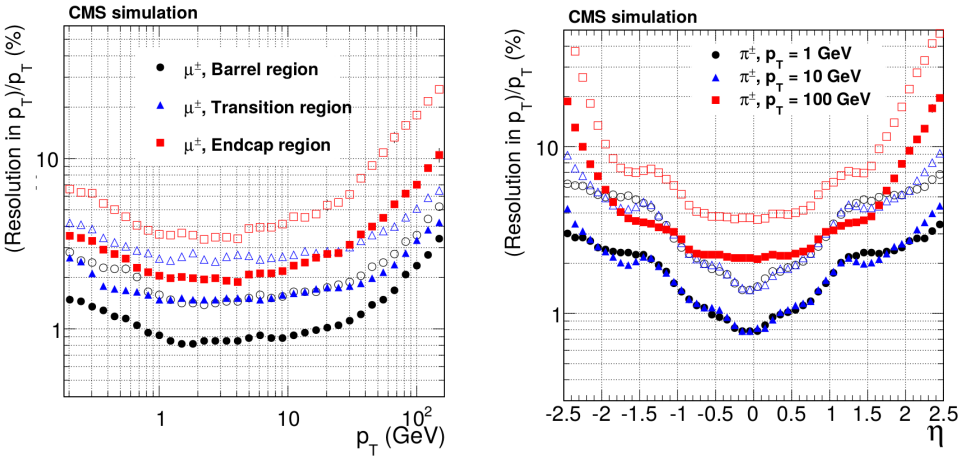


Figure 5.2: Relative transverse momentum resolution as a function of muon p_{T} (left) and pion pseudorapidity (right). For each bin in p_{T} and η, the solid (open) symbols correspond to the interval covering the 68% (90%) of the residual $(p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}})$ distribution. Figures are taken from [65].

High-purity tracks

High-purity (HP) track category is a collection of tracks with stringent selection requirements and low fake rate. The HP tracks are selected based on the number of layers with at least one hit, χ^2/dof of the track fit, and a maximum number of tracker layers without hits. In addition, the transverse and longitudinal impact parameter significance requirements are imposed. The tracks reconstructed with different iteration rounds have different selection requirements for the high-purity collection. More detailed description of the high-purity selection can be found in Table 5 of Reference [65].

5.2 Vertex reconstruction

Several charged particles are created in the collision of two protons. The reconstruction of the tracks helps to obtain the position of the original interaction point of the protons, which are referred to as primary vertices. On the other hand relatively long-lived particles, such as B mesons, may decay rather far away from the beam line. The decay positions are called secondary vertices. Precise and careful reconstruction of the primary and secondary vertices is important for B physics analyses, which utilise the proper decay time measurement of the b hadron candidates.

5.2.1 Pixel vertices

A fast reconstruction of pixel vertices [65] is the first step in track reconstruction. The track stubs are formed of three hits in different pixel layers and have $p_{\mathrm{T}} > 0.9$ GeV. The tracks are then assigned to the vertices using a simple gap clustering algorithm. The algorithm orders the track stubs in descending order with respect to the distance of closest approach from the beam spot. If any of the two neighbouring tracks have a separation of $z = 2$ mm, the tracks on either side of the cut are separated into two vertices. The vertices are created by fitting the tracks assigned to the vertices with an adaptive vertex fit [66]. The adaptive vertex fit is an iterative fit that down-weights the tracks according to increasing distance from the track cluster. The weights are varied from one iteration to another until the convergence is reached.

5.2.2 Primary and secondary vertices

The primary vertex reconstruction [65] proceeds in three phases. First the tracks originating from the beam spot are identified based on a set of selection criteria. The beam spot is the luminous region of the detector where two beams of protons

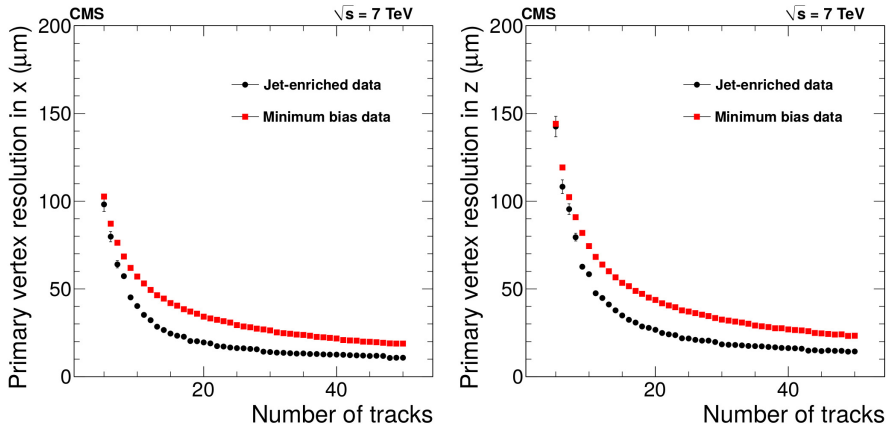


Figure 5.3: Position resolution of the primary vertices as a function of the number of fitted tracks. The resolution in x-coordinate is shown on the left and the resolution of z-coordinate is shown on the right. The resolution in y is similar to that of the x-coordinate and is not presented here. Figure is taken from [65].

collide. The track selection involves requirements for value of the transverse impact parameter significance, numbers of pixel and strip hits associated with the tracks, and normalised χ^2 from the track fit.

The selected trajectories are then clustered using the z-coordinates at their points of closest approach to the center of the beam spot. The algorithm used in the clustering is deterministic annealing [67]. Finally, an adaptive vertex fit is performed to obtain the most precise vertex position coordinates in three dimensions.

The position resolution of the primary vertices depends on the number of tracks and the average track p_{T}. Resolutions of x- and z-coordinates for minimum bias tracks and jet-enriched data with transverse energy $E_T > 20$ GeV are shown in Figure 5.3. The primary vertex resolution for the minimum bias tracks is about 25 and 30 μm in x- and z-coordinates with vertices of at least 40 tracks. The resolutions for the jet-enriched data are around 15 μm for both x- and z-axes. The resolution in y is similar to that of the x-coordinate and thus it is not shown in here.

The secondary vertices can be reconstructed with the Kalman filter technique,

but the adaptive vertex fitter can also be used. Both methods result in similar secondary vertex resolutions. The number of tracks originating from the secondary vertex is much smaller compared to the number of tracks arising from the primary vertices. For instance, the $B_s^0 \to J/\psi\phi$ decay vertex has four out-coming tracks, two muons and two kaons, while tens of tracks originate from primary vertices. The small number of tracks leads to a lower position resolution for the secondary vertices compared to the resolution determined for the primary vertices. The secondary vertex resolution for $B_s^0 \to J/\psi\phi$ process is predicted to be about 55 μm in transverse plane and about 75 μm for the z-coordinate [1]. The resolution estimates are obtained with simulated $B_s^0 \to J/\psi\phi$ events.

5.3 Beam spot

The center of the beam spot (BS) and its uncertainty are determined in each lumisection that corresponds to 23 seconds of data taking. The position coordinates of the BS can be determined using two methods [65]. The mean BS position can be fitted to the 3D distribution of the reconstructed primary vertex positions. The second method takes advantage of correlations between track d_0 and ϕ coordinates. The track d_0 can be parametrised as a function of ϕ, z_0 and the BS coordinates, and a fit can be performed to the distribution of track parameters. The results given by the two fits are in agreement. Both fits are used in the CMS reconstruction software. The $d_0 - \phi$ fit gives the BS coordinates in the transverse plane whereas the 3D vertex position fit determines the z-coordinate and the uncertainties of the BS position.

The x and y position uncertainties of the BS location can also be obtained with an alternative fit [65]. The method is based on event-by-event correlations between the d_0 parameters of two tracks that arise from the same vertex. The strength of the correlation is proportional to the transverse size of the beam. The

uncertainties given by alternative method are found to be consistent with the 3D primary vertex fit. The beam spot position varies from lumisection to lumisection, but the average uncertainty of the BS coordinates in transverse plane is within 2-3 μm. The comparison of the vertex resolutions for the primary and secondary vertex, as well as for the beam spot in the transverse plane, is shown in Table 5.1.

	$\sigma_{x,y}$ (μm)
Primary vertex	15
Secondary vertex	55
Beam spot	2-3

Table 5.1: Comparison of vertex and beam spot resolutions in transverse plane. The primary vertex resolution is determined with the vertices that contain 40 tracks. The secondary vertex resolution is measured using simulated $B_s^0 \to J/\psi\phi$ decays.

5.4 Muons

Muon tracks are reconstructed from the hits in the tracker and in the muon chambers. The muon reconstruction [1] starts from the search of the track segments in the drift tubes and cathode strip chambers that are embedded in the return yoke. The segments are then connected to so called stand-alone muon tracks proceeding from one muon station to another. Depending on the reconstruction strategy, a tracker track can be extrapolated to the muon spectrometer, or a stand-alone muon can be propagated backwards to the tracker. The two reconstruction techniques are referred to as tracker and global muons, respectively.

The track segments in the drift tubes are approximated as straight lines. The segments are reconstructed independently in the transverse ($r-\phi$) and longitudinal ($r-z$) planes. The two views are then combined in to a 4-dimensional segment that includes two position coordinates and two angles.

In the CSCs, hit positions are formed in the transverse plane by combining hit information from the anode wires and the cathode strips. A CSC track segment is first reconstructed using two hits, one from the first and another from the last of the six layers of a CSC chamber and approximating a straight line between the hits. Then, an attempt is made to associate more hits on the segment by investigating the compatibility of the hits in intermediate layers and updating the linear fit respectively. The segment must be reconstructed from at least four hits, otherwise it is discarded.

The stand-alone muon tracks in the DT chambers are formed by searching for compatible segments in different stations with the Kalman filter technique. In the CSC chambers, three-dimensional hits associated with the muon segments are exposed to the Kalman filter fit. The hits in the RPCs are also included in the track finding procedure. The track propagation takes into account multiple scattering, energy loss, and effects of a non-uniform magnetic field.

5.4.1 Global muon reconstruction

Global muons [68] are reconstructed starting from a trajectory of stand-alone muon and searching for compatible hits in the tracker using the Kalman filter fit. The global fit can improve the momentum resolution with respect to the tracker-only fit for those tracks having transverse momentum order of 200 GeV or larger.

5.4.2 Tracker muon reconstruction

Tracker muon reconstruction [68] starts by considering all the tracks with $p_{\mathrm{T}} > 0.5$ GeV and $p > 2.5$ GeV as muon candidates. The candidate tracks are then extrapolated to the muon spectrometer by taking into account the magnetic field, the average track energy loss, and the uncertainty due to multiple Coulomb scattering in the detector material. If at least one muon segment in CSCs or DTs matches

with the tracker track, the track is classified as a tracker muon. The tracker muons improve reconstruction efficiency for muons having $p_{\mathrm{T}} < 5$ GeV.

5.4.3 Muon categorisation

There are several sets of muon selection criteria optimised for different types of physics analyses. Two common muon categories used in the B physics analyses are soft muons and loose muons. The muons reconstructed with the particle flow (PF) algorithm [69] are also utilised in the B physics measurements.

Particle flow muons

Particle flow muons [69] are tracker or global muons reconstructed with an efficient particle identification technique called particle flow. The PF algorithm uses information from several subdetectors to identify the electrons, muons, photons, and neutral and charged hadrons in the collision event. The PF algorithm is optimised for efficiently identifying the muons within jets while retaining low misidentification rate for charged hadrons. The PF algorithm is discussed with more detail in Section 5.6.

Soft and loose muons

High-purity tracker tracks that are matched to a muon segment are classified as soft muons [70]. The requirement for the matching is that the pulls for the x- and y-coordinates of the segment and the track are less than three. Stringent requirements are also applied for the tracker track. The track selections are listed in Table 5.2.

Loose muons are tracker or global muons identified with the PF algorithm. In contrast to the soft muons, loose muons do not have various selection criteria to the track quality. For the analyses containing multiple muons, a muon pair must satisfy the opening angle requirement $\Delta\mathrm{R} > 0.02$. The cut is applied in order to

exclude fake muon pairs that in reality originate from a single trajectory split into two during the muon segment reconstruction [70].

The reconstruction and identification efficiencies of soft and PF muons are shown in Figure 5.4. Both selections are fully efficient when muon transverse momentum is larger than 3 (4) GeV. The threshold value, $p_{\mathrm{T}} > 3\,(4)$ GeV, depends on the muon pseudorapidity.

Number of tracker layers with hits ≥ 6
Number of pixel layers ≥ 1
Transverse impact parameter $d_{xy} < 0.3$ cm w.r.t the primary vertex
Longitudinal impact parameter $d_z < 20$ cm w.r.t. the primary vertex

Table 5.2: Selection requirements of the soft muon category [70].

5.5 Electrons

Electrons can lose a significant amount of their energy via brehmstrahlung. About 35% of electrons radiate more than 70% of their total energy before entering electromagnetic calorimeter [71]. This has a significant effect on the curvature of the electron trajectory, which must be accounted for in the track reconstruction.

There are two ways to identify electrons. One can begin the reconstruction from the ECAL clusters [71], or start from tracker tracks using the particle flow approach. The PF algorithm is discussed in Section 5.6 while the other reconstruction technique is described here.

Electron reconstruction [71] begins by identifying the energy deposits of the brehmstrahlung photons and collecting them in ECAL superclusters. Clustering is performed with two algorithms [72]. The hybrid algorithm is used in the ECAL barrel region while the island algorithm is applied in the endcaps.

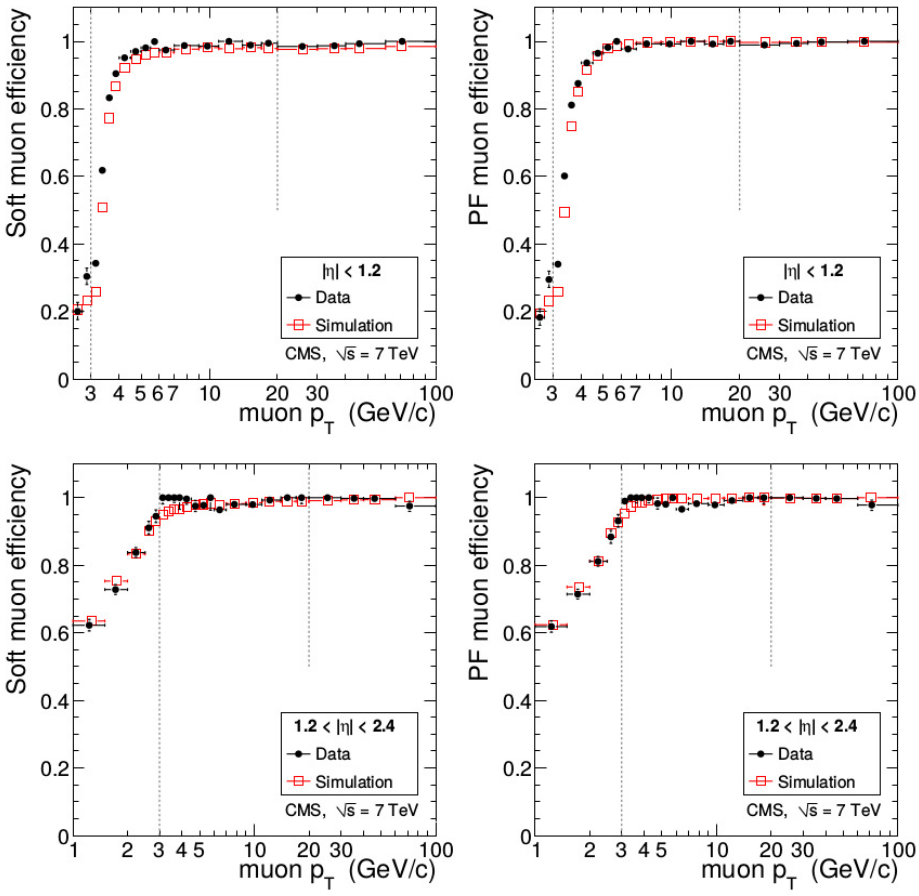


Figure 5.4: Muon reconstruction and identification efficiency for soft and PF muons as a function of muon p_{T} given that the tracker track exits. The efficiencies are measured from $\mathrm{J}/\psi \to \mu\mu$ and $Z \to \mu\mu$ events. The $\mathrm{J}/\psi \to \mu\mu$ decay is used for the muons with $p_{\mathrm{T}} < 20$ GeV and while the Z-boson decay are used for the muons with $p_{\mathrm{T}} \geq 20$ GeV. Figure is taken from [68].

The track seeds of electrons are reconstructed with the different approach compared to the nominal one covered in Section 5.1. The electron track seeds are identified by extrapolating the energy-weighted mean position of the superclusters backwards to the pixel tracker. A hit compatible with the position of the supercluster is first searched for the innermost pixel layer within a loose $\Delta\phi$, Δz window. If no

suitable hit is found in the first pixel layer, the hit is looked for in the next layer. In either case, the track is then propagated to the following layer(s) of the pixel tracker.

The Kalman fitter cannot be used to fit the electron tracks as it assumes a Gaussian distribution for the energy deposits of the hits in the tracker. Due to the brehmstrahlung emission the energy deposits of the hits have a non-Gaussian distribution. To account for the effect, the track fit is performed with a Gaussian sum filter (GSF) algorithm. When the pixel track is reconstructed, the GSF algorithm looks for compatible hits in the strip tracker. The track is extrapolated to the next strip layers using the Bethe-Heitler model of the electron energy loss. The track propagation is terminated if the last strip layer is reached or if there are two consecutive layers without a suitable hit. At least five tracker hits are required to construct an electron track.

5.6 Particle flow algorithm

The particle flow algorithm [69] combines information from all CMS subdetectors. It identifies the particles created in the proton-proton collisions: electrons, muons, and photons as well as neutral and charged hadrons. Jets are also reconstructed from the particles identified by the PF algorithm. Missing transverse momentum is determined by adding up the transverse momenta of the PF candidates.

The core elements of the particle flow algorithm are the calorimeter clusters, tracker tracks, and muon segments that are reconstructed with the standard CMS software. A link algorithm then creates links between the ECAL clusters and tracker trajectories, muon segments and tracker tracks, and ECAL and HCAL clusters, as well as tracks and calorimeter clusters. The elements linked to each other form blocks.

Using the blocks of linked elements, the PF algorithm reconstructs and iden-

tifies particles. The PF muons are formed from global muons if the momentum of the global muon is within three standard deviations of the momentum determined from the tracker track. When the PF muon is identified, its track is removed from the block.

Electrons are reconstructed and identified by examining the tracker tracks in each block. If the tracks fulfil the criteria of the typical electron track, the trajectory is fitted with the GSF algorithm to the border of the strip tracker and ECAL. Final electron identification is carried out later by using a number of tracking and calorimeter variables.

The remaining tracker tracks are filtered by requiring higher p_{T} resolution in the tracker compared to the energy resolution in the calorimeters. The total momentum of the tracks linked to ECAL and HCAL clusters are compared to the total energy deposited to the calorimeters. Energy deposits in the calorimeters are required to be smaller than the energy of the tracks. If the energy deposits in the ECAL and HCAL are significantly lower than the energy of the tracks, an additional muon search is initiated, and the muon energy deposits 3 (0.5) GeV in ECAL (HCAL) are removed from the calorimeter clusters. The remaining tracks in the block are identified as particle flow charged hadrons whose momenta and energies are determined directly by the track under a charged pion mass hypothesis.

If the energy of the closest calorimeter clusters linked to the tracks is larger than the energy of tracks in the tracker, the energy excess is cast to the either photons or neutral hadrons. The energy deposits of remaining ECAL and HCAL clusters not linked to any tracks are also assigned to photons and neutral hadrons.

5.7 Composite particles

Kinematic observables and the vertex position of a composite particle such as a $\mathrm{B_s^0}$ meson decaying into $\mathrm{J}/\psi\phi$ state can be reconstructed with a kinematic fit technique

[73] using the muon and kaon tracks from $J/\psi \to \mu\mu$ and $\phi \to KK$ processes. The algorithm is designed to find the most optimal vertex position (x, y, z), momentum $\boldsymbol{p}$ and mass m of a mother particle, assuming that the N final state particles with parameters (x_n, y_n, z_n), $\boldsymbol{p}_n$ and m_n originate from it.

The algorithm uses the least mean squared minimization to find the optimal values for the seven parameters of the composite particle. The minimization can also be constrained utilising information from the decay process. For instance, the muons originate from the J/ψ meson, whose mass lies in a certain mass range and the muon and kaon tracks originate from a common vertex. The mathematical details of the minimization problem can be found in References [73, 74].

Two approaches of the kinematic fit can be used. A global fit strategy determines the unknown parameters in the single fit by refitting the input data with all the constraints. A sequential approach exploits series of individual fits where the constraints are applied one-by-one. The sequential strategy becomes is useful when the decay chain contains long-lived unstable particles that traverse through significant distances in the detector. Mathematically both approaches are equivalent. The global fit approach is exploited in the reconstruction of the B mesons described in the following chapters.

The kinematic fit with mass constraints of the final state particles improves the mass resolution of the composite particles compared to the Kalman fit technique. Studies with simulated samples show that the B_s^0 mass resolution obtained from the kinematic fit is by a factor of two better compared to the resolution achieved with the Kalman filter technique [1]. However, a constrained kinematic fit does not affect the decay vertex resolution.

6. Flavour tagging in the CMS experiment

Two flavour tagging algorithms for the weak mixing phase measurement are described in this chapter. The first one is a cut-based algorithm [75] that established the CMS flavour tagging framework and tagging guidelines for the more complex multivariate tagger. This algorithm is used in the early weak mixing phase measurement [11]. The more advanced algorithm [76] using a multilayer perceptron (MLP) neural network (NN) is exploited in the latest weak phase measurement [12] of the CMS experiment.

6.1 Data and simulated samples

The data used in the flavour tagging studies were collected in proton-proton collisions during the 2012 run of the LHC at a centre-of-mass energy of 8 TeV, and correspond to an integrated luminosity of 19.7 fb^{-1}. Simulated Monte Carlo (MC) samples $\text{B}_\text{s}^0 \to \text{J}/\psi(\to \mu\mu)\phi(\to \text{K}^+\text{K}^-)$, $\text{B}^+ \to \text{J}/\psi(\to \mu\mu)\text{K}^+$ and $\text{B}_\text{d}^0 \to \text{J}/\psi(\to \mu\mu)\text{K}^{0^*}(\to \text{K}^+\pi^-)$ are utilised in the development and validation of the flavour tagging algorithms.

The events in the MC samples are generated with PYTHIA 6.4 [77], EvtGen [78], PHOTOS [79] and GEANT4 [80] packages. The hard parton interaction of the colliding protons is simulated with PYTHIA 6.4 that uses parton distribution functions to obtain the energies of the partons in the interaction. The transition

amplitudes of the interesting physics processes are evaluated at the leading order of perturbative QCD. When the transition amplitudes are determined, the parton shower algorithm simulates the complex 2 → n interactions applying the string model [81]. In addition to the partons participating in the hard interaction, several soft quark and gluon interactions are also simulated.

The decays of short-lived particles such as b and c hadrons are simulated with the EvtGen generator. Final state radiation is included in the EvtGen simulation using the PHOTOS package. Propagation of long-lived particles through the detector volume, their interactions with detector material, and the sensor responses are managed with the GEANT4 package.

6.2 Tagging method

Flavour tagging is performed to simulated $B_s^0 \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ decays and to B^+ events reconstructed from the data. Simulated B_s^0 events are used to optimise the two algorithms. The tagging performances obtained from the simulated B_s^0 and B^+ decays are compared to check the flavour independence of the tagging results. The neural network-based algorithm also uses simulated $B_d^0 \to J/\psi K^{0*}$ decays for testing and validation of the tagging performance. The final tagging performance is determined using $B^+ \to J/\psi K^+$ data. In addition, the performance of the neural network-based algorithm is verified with a measurement of B_d^0 mixing asymmetry using the $B_d^0 \to J/\psi K^{0*}$ events reconstructed from the data.

6.3 Event reconstruction and selection

Reconstruction of B mesons is based on the identification of the $J/\psi \to \mu\mu$ decay together with the reconstruction of a $\phi(1020) \to KK$ candidate, a charged track (K^+), or a $K^*(892)^0 \to K\pi$ candidate, depending on the decay mode. The optimisa-

tion of the $B_s^0 \to J/\psi\phi$ event selection is presented in Reference [82]. The selection was first used in the measurement of B_s^0 production cross section [83]. The selection requirements of B^+ and B_d^0 events are chosen to be as similar as possible to those applied to $B_s^0 \to J/\psi\phi$ decays.

6.3.1 Trigger selection

The data are collected with a high-level trigger optimised for detecting the $J/\psi \to \mu\mu$ decay that is displaced from the beam spot. The trigger reconstructs a J/ψ meson from an oppositely-charged muon pair using a set of selection criteria. The muons are required to have $p_T > 4$ GeV and $|\eta| < 2.2$. The total transverse momentum of the muon pair has to be greater than 6.9 GeV.

Two muon tracks are fitted to a common vertex. The minimum χ^2 vertex fit probability is required to be 10%. A transverse decay length significance $L_{xy}/\sigma_{L_{xy}}$ has to be greater than three, where L_{xy} is the transverse distance between the beam spot and the secondary vertex, and $\sigma_{L_{xy}}$ is its uncertainty. The muon tracks are required to have the distance of closest approach (DCA) less than 0.5 cm. The J/ψ mass is required to be within 2.9-3.3 GeV window. The cosine of the opening angle between the transverse decay length vector $\boldsymbol{L}_{xy}$ and the J/ψ transverse momentum vector has to satisfy the $\cos(\boldsymbol{L}_{xy}, \boldsymbol{p}_T) > 0.9$ requirement.

6.3.2 J/ψ mesons

The J/ψ mesons are reconstructed from oppositely charged muon pairs. The muons have to be classified as PF muons and matched to the muons detected by the trigger. The muons are required to originate from a common vertex whose position and validity is determined using the Kalman vertex fit. The four-momentum of the J/ψ candidates are obtained by summing the four-momenta of the muon tracks. The mass of the J/ψ candidate is required to be within 150 MeV of the world-average

J/ψ mass [84]. In order to have a consistent selection for the data, the other offline cuts for the muon properties such as p_{T} and DCA and J/ψ properties such as p_{T} and vertex probability are set at least as tightly as in the trigger selection.

6.3.3 $\phi(1020)$ mesons

The $\phi(1020) \rightarrow \mathrm{KK}$ candidates are reconstructed from oppositely charged tracks that have $p_{\mathrm{T}} > 0.7$ GeV. Each track has to comprise at least five tracker hits. The tracks are selected from a collection where the muon candidate tracks have been removed. Tracks are assumed to be kaons, as no particle identification is available for the tracks. The invariant mass of the track pair is required to lie within 10 MeV of the world-average $\phi(1020)$ mass [84].

6.3.4 $\mathrm{K}^*(892)^0$ mesons

The $\mathrm{K}^*(892)^0 \rightarrow \pi^-\mathrm{K}^+$ mesons are reconstructed using the same track collection and track selection requirements as used in the ϕ meson reconstruction. Because of the lack of particle identification detectors in the CMS apparatus, the kaons and pions cannot be reliably differentiated. Hence, the tracks always have two mass hypotheses. The tracks are first assigned to be π^- and K^+ ($\mathrm{K}^*(892)^0$) and then K^- and π^+ ($\overline{\mathrm{K}}^*(892)^0$). The mass of the $\mathrm{K}^*(892)^0$ ($\overline{\mathrm{K}}^*(892)^0$) candidate must to lie within 100 MeV of the world-average mass [84]. The assignment that results in a $\mathrm{K}^*(892)^0$ ($\overline{\mathrm{K}}^*(892)^0$) mass closest to the world-average is retained while the other mass hypothesis is rejected. However, there are always a number of $\mathrm{K}^*(892)^0$ ($\overline{\mathrm{K}}^*(892)^0$) candidates having misassigned kaon and pion masses since the meson has a large decay width of about 50 MeV [84].

6.3.5 B_s^0 mesons

The B_s^0 mesons are formed by performing a kinematic fit to the muon and kaon tracks that originate from the J/ψ and $\phi(1020)$ candidates, respectively. The dimuon mass is constrained to the nominal J/ψ mass [84] in a kinematic fit. A B_s^0 candidate is retained if its invariant mass lies within the window of 5.20-5.65 GeV and its χ^2 vertex fit probability is greater than 2%. If the event has multiple B_s^0 candidates, the one having the highest vertex probability is selected.

6.3.6 B^+ and B_d^0 mesons

The B^+ and B_d^0 candidates are formed by using a kinematic fit to the final state tracks and constraining the dimuon mass to the world-average J/ψ mass. The kaon track used to reconstruct the B^+ candidate is also required to have p_T greater than 2 GeV. The tracks originating from the decays of $K^*(892)^0$ candidates must have a minimum transverse momentum of 0.7 GeV. The masses of the B^+ and B_d^0 candidates are required to lie within the range of 5.05-5.55 GeV and 5.18-5.43 GeV, respectively, and their χ^2 vertex fit probabilities have to be greater than 2%. If the event contains more than one reconstructed B^+ (B_d^0) meson, the candidate having the highest vertex probability is retained.

6.4 B^+ mass fit

The numbers of mistagged, tagged, and signal events in the $B^+ \to J/\psi K^+$ data sample are determined using an unbinned extended maximum-likelihood fit of the model probability density function (pdf) to the B^+ mass distribution. The fit is applied to the mistagged, tagged and signal events simultaneously. The shape of the signal mass distribution is assumed to be independent of the tag classification. All the pdf parameters are the same in the fitted samples except for the parameters

that characterise numbers of events in each sample.

The signal component of the pdf is described by a double Gaussian function with a common mean value. The background is modelled with a sum of the first order Chebyshev polynomial and a modified error function. The likelihood function is defined as

$$L_{tot}(m) = N_{sig}L_{signal}(m) + N_{bkg}L_{background}(m),$$

where $L_{signal}(m)$ and $L_{background}(m)$ are the signal and background functions respectively and N_{sig} and N_{bkg} are the numbers of signal and background events. The signal and background likelihood functions are

$$\begin{aligned} L_{signal}(m) &= \prod_{i=1}^{N}\sum_{j=1}^{2} f_j \frac{1}{\sqrt{2\pi}\sigma_{m,j}} e^{-(m_i-\mu_m)^2/2\sigma_{m,j}} \\ L_{background}(m) &= \prod_{i=1}^{N} f_{BG}\cdot(km_i) + (1-f_{BG})\cdot\{\mathrm{Erf}\left(\frac{-m_i+5.15}{a}\right)+1\}, \end{aligned}$$

where

- the signal mass function is a sum of two Gaussians with a common mean (μ_m) and widths $\sigma_{m,i}$. The parameters f_j are the relative strength between the Gaussians and they add up to unity $\sum_{j=1}^{2} f_j = 1$.
- Function km_i is the first-order Chebyshev polynomial with slope k.
- $\mathrm{Erf}((-m_i+5.15)/a) = \frac{2}{\sqrt{\pi}}\int_0^{(-m_i+5.15)/a} e^{-x^2}dx$ is a modified error function and a is a constant coefficient, $a = 0.0324$. Parameter f_{BG} is the relative strength between the two background functions, $0 < f_{BG} < 1$.

6.5 Simulation bias

The tagging performance of the simulated samples is affected by the b hadron composition in the opposite side (OS) of the $b\bar{b}$ production. For instance, if the opposite

side b ($\bar{\mathrm{b}}$) quark hadronized always to the $\bar{\mathrm{B}}_{\mathrm{s}}^{0}$ ($\mathrm{B}_{\mathrm{s}}^{0}$) meson, the fast flavour oscillation would result in practically zero tagging power. The opposite-side b hadron fractions for two different simulated samples and the world-average values of the production fractions [2] are shown in Table 6.1. The quantities f_s, f_u, f_d and f_{baryon} refer to the production fractions of the $\mathrm{B}_{\mathrm{s}}^{0}$, B^{+}, $\mathrm{B}_{\mathrm{d}}^{0}$ mesons and b baryons (mostly Λ_{b}), respectively. A small fraction of $\mathrm{B}_{\mathrm{c}}^{+}$ mesons ($\sim 0.03\%$) on the opposite side of the MC samples and in data ($\sim 0.2\%$) [2] is omitted.

	$\mathrm{B}_{\mathrm{s}}^{0} \to \mathrm{J}/\psi\phi$ MC	$\mathrm{B}^{+} \to \mathrm{J}/\psi\mathrm{K}^{+}$ MC	world-average (2016)
f_s [%]	6.8 ± 0.1	15.1 ± 0.1	10.0 ± 0.8
f_u [%]	42.0 ± 0.1	25.2 ± 0.1	40.7 ± 0.7
f_d [%]	42.1 ± 0.1	49.1 ± 0.1	40.7 ± 0.7
f_{baryon} [%]	9.1 ± 0.1	10.6 ± 0.1	8.5 ± 1.1

Table 6.1: The production fractions of the opposite-side b hadrons found from the simulated generator-level $\mathrm{B}_{\mathrm{s}}^{0} \to \mathrm{J}/\psi\phi$ and $\mathrm{B}^{+} \to \mathrm{J}/\psi\mathrm{K}^{+}$ samples together with the world-average [2] production fractions. A small fraction of $\mathrm{B}_{\mathrm{c}}^{+}$ mesons on opposite-side of the $\mathrm{b}\bar{\mathrm{b}}$ hadronization is omitted.

The signal selection and reconstruction procedure discards the events that do not contain the signal hadron e.g. a $\mathrm{B}_{\mathrm{s}}^{0}$ or a B^{+} meson. This makes the probability to observe a pair of two signal particles, for instance a ($\mathrm{B}_{\mathrm{s}}^{0}$, $\bar{\mathrm{B}}_{\mathrm{s}}^{0}$) pair, different from the one obtained in the generic $\mathrm{b}\bar{\mathrm{b}}$ production and hence affects the overall opposite-side b hadron composition. Therefore, the tagging results determined from the simulated $\mathrm{B}_{\mathrm{s}}^{0}$ and B^{+} samples have to be re-weighted in order to make a meaningful comparison of tagging performances. The origin of the bias in the opposite-side b hadron composition and the re-weighting procedure were first published in Reference [76]. A description of the bias and re-weighting method is also given below.

The fraction of events containing at least one signal B_q meson is

$$f_q(1-f_q) + (1-f_q)f_q + f_q^2 = 2f_q(1-f_q) + f_q^2 = f_q(2-f_q) \tag{6.1}$$

where the terms $f_q(1-f_q)$ represent the fraction of events with exactly one signal B_q meson and f_q^2 is the fraction of events containing a (B_q,B_q) pair. The biased abundance of signal B_q hadrons on the opposite side becomes

$$f_q' = \frac{f_q^2}{f_q(2-f_q)} = \frac{f_q}{2-f_q} < f_q. \tag{6.2}$$

However, Equation (6.2) does not account for the effect of selection and reconstruction efficiency ϵ_q which affects the data and simulated samples differently. When both b quarks hadronize into B_q, the probability for reconstructing at least one B_q meson is given by $1-(1-\epsilon_q)^2 = \epsilon_q(2-\epsilon_q)$. The modified OS fraction can then be written as

$$f_q' = \frac{f_q^2(1-(1-\epsilon_q)^2)}{2\epsilon_q f_q(1-f_q) + (1-(1-\epsilon_q)^2)f_q^2} = \frac{f_q(2-\epsilon_q)}{2-f_q\epsilon_q}. \tag{6.3}$$

For the $B_s^0 \to J/\psi\phi$ ($B^+ \to J/\psi K^+$) events in the data, the efficiency including detector acceptance, trigger, reconstruction and selection efficiencies is rather low, $\epsilon_q << 1$. Under an approximation of $\epsilon_q \approx 0$, the fraction f_q' coincides with the production fraction: $f_q' \approx f_q$.

In the simulated $B_s^0 \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ samples at least one B_s^0 (B^+) meson is always produced within the detector acceptance in each generated event. This leads to a large selection efficiency that can be evaluated to be $\epsilon_q \approx 1$. Hence, for the simulated events a high selection efficiency causes a bias in the opposite side b hadron fractions and $f_q' \approx \frac{f_q}{2-f_q}$. A correction factor w_q is determined and applied to the events with exactly two generated signal B_q mesons. The weight w_q is obtained from the equation

$$f_q = \frac{w_q f_q'}{w_q f_q' + f_o'}, \tag{6.4}$$

where $f'_o = 1 - f'_q$. By substituting f'_q using Equation (6.3) and solving for w_q, one obtains

$$w_q = \frac{2}{2 - \epsilon_q}. \tag{6.5}$$

For the simulated samples $w_q \approx 2$ as the selection efficiency $\epsilon_q \approx 1$, while for the data $w_q \approx 1$, since $\epsilon_q \approx 0$. Therefore, no correction is applied to the data, whereas $w_q = 2$ is used to correct for the bias in the opposite-side b hadron compositions in the simulated samples.

6.6 Cut-based tagging algorithm

The design of the tagging algorithm aims at identifying electrons and muons that arise from the decays of opposite-side b hadrons. As already discussed earlier, the charge of the lepton is correlated with the flavour of the B signal meson. The algorithm cuts on the lepton observables, which enhance the selection of signal leptons and reject the background leptons.

The tag muons used in the design of the tagging tool are identified with the PF algorithm, while the tag electrons are reconstructed with a cluster-based method and called Gaussian-sum-filter (GSF) electrons. The muons that originate from the decays of the signal side J/ψ meson are not used in tagging. The electron tracks overlapping with muon tracks are also omitted from the electron collection in which the tag electrons are searched. If an event contains multiple electrons or muons suitable for tagging, the lepton that has the highest p_T is selected to tag the B_s^0 meson.

The tagging performance of baseline taggers without any selection are shown in Table 6.2. Although the tagging efficiency is high, the mistag fractions (ω) for tagged electron and muons are near 50%, which means that there is hardly any correlation between the charge of the lepton and the flavour of the signal meson.

Thus, optimisation of the lepton selection has to be performed before using the flavour tagging algorithm in the weak mixing phase measurement.

	Muons	Electrons
Mistag fraction ω [%]	45.3 ± 0.3	47.7 ± 0.2
Tagging efficiency ϵ_{tag} [%]	20.8 ± 0.1	25.2 ± 0.1
Tagging power P_{tag} [%]	0.19 ± 0.02	0.05 ± 0.01

Table 6.2: A baseline tagging performances of electrons and muons without optimisation. The tagging performances are measured from simulated $B_s^0 \to J/\psi\phi$ sample.

6.6.1 Input variables

The input observables used in the design of the tagging algorithm are i) lepton transverse momentum, ii) impact parameter of the lepton with respect to the primary vertex associated to the signal B meson, and iii) angular distance between the lepton and the signal B meson $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. An additional cut on multivariate electron identification discriminator is imposed for the tag electrons. The variables are mostly used to discriminate between the leptons coming from the semileptonic b-hadron decays and the leptons arising from the background processes. The distributions of the input observables from the simulated decays and from the data are consistent and are shown in Appendix A.

- Lepton transverse momentum

  The transverse momentum of the lepton is useful for rejecting the leptons originating from background processes such as decays of light hadrons. The leptons arising from the decays of background hadrons are often characterised by lower p_T compared to those from $b \to lX$ decays.

- Lepton impact parameter (IP)

As the b hadrons have lifetimes of about 1.5 ps, the leptons created in the semileptonic decays may have a relatively large distance of closest approach from the primary vertex. On the other hand, leptons that have a very large impact parameter are more likely to arise from other primary vertices and are not correlated with the signal b hadron decay. Thus, the impact parameter is plays an important role for both the flavour identification and background rejection. The impact parameter is defined such that the lepton track is first extrapolated to the point of closest approach with respect to the beam axis named as impact point. The impact parameter of a lepton is calculated as a three-dimensional distance between the primary vertex and the impact point of the track.

- Angular distance between the lepton and the signal b hadron (ΔR)

  The angular distance between the lepton and the b hadron is used to reject the particles originating from the signal-side of $b\bar{b}$ hadronization. A charged kaon can be created on the same side as the signal-B meson. If the kaon produced in the hadronization process of the signal meson is misidentified as an opposite-side lepton, the reversed charge-flavour correlation will result in a mistagged event.

- Electron identification variable

  A cut for a multivariate discriminator [85] is imposed for tag electrons to decrease the misidentification rate of non-triggered electrons.

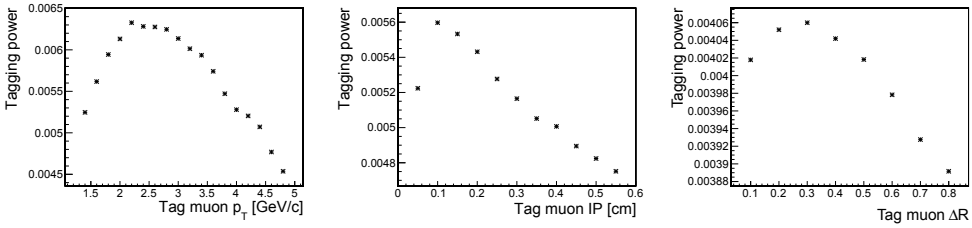
6.6.2 Optimisation

The tag leptons are identified with selection requirements set on the lepton input observables. These requirements are determined in order to maximise the tagging power $P_{tag} = \epsilon_{tag}(1 - 2\omega)^2$ in the simulated B_s^0 sample. The optimisation is per-

formed on muon and electron samples separately. The optimised lepton selection requirements are shown in Table 6.3. The measured tagging power as a function of the cuts on the muon and electron tagging variables are shown in Figures 6.1 and 6.2. Only one observable is varied at the time, while no cuts are set on the other variables.

	electron	muon
impact parameter (IP)	< 1.0 mm	< 1.0 mm
transverse momentum (p_{T})	> 2.0 GeV	> 2.2 GeV
angular distance ΔR	> 0.2	> 0.3
electron MVA-discriminator	> -0.6	-

Table 6.3: Selection requirements for the tag leptons.



(a) **(b)** **(c)**

Figure 6.1: Tagging power as a function of tag muon p_{T}, IP and ΔR.

6.6.3 Tagging performance of the cut-based algorithm

The average tagging performances for the electron and muon taggers are determined separately. Then the mistag fractions of the both taggers are binned and finally parametrised as a function of the lepton p_{T}. A combined average tagging performance is obtained by merging the two taggers in a single class of tag leptons.

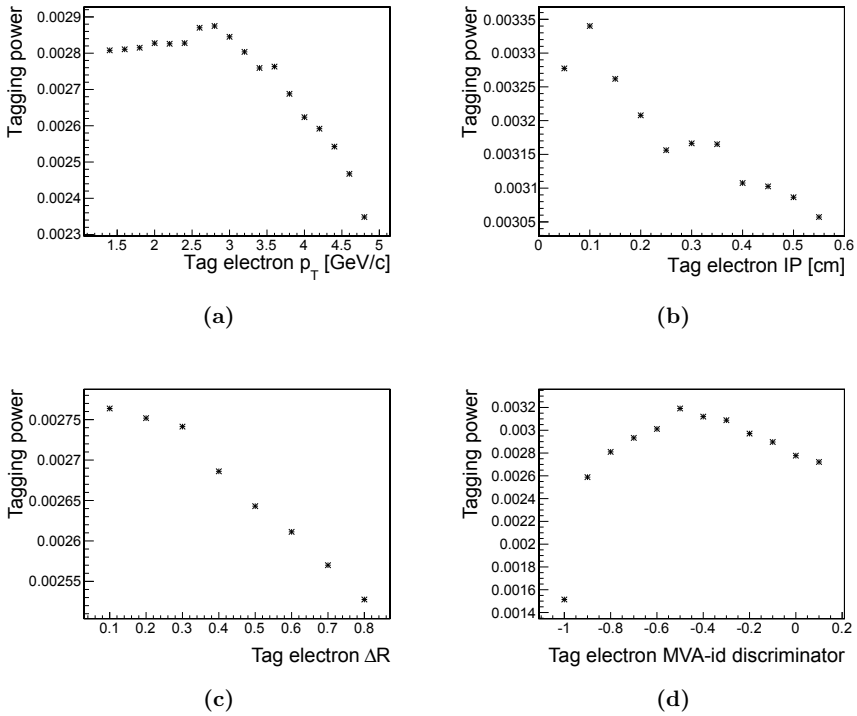


Figure 6.2: Tagging power as a function of tag electron p_{T}, IP and ΔR as well as MVA-discriminator.

If an event is tagged with both lepton species, the lepton that has a higher absolute value of dilution, $|D| = |1 - 2\omega|$, is selected to tag the event. On the other hand, if the mistag fraction associated with the lepton is larger than 0.5, the hypothesis for the B meson flavour is reverted and the mistag fraction is changed accordingly, $\omega' = 1 - \omega$.

Average tagging performance for electrons and muons

The average tagging performances for muons and electrons are shown in Tables 6.4 and 6.5. The yields of wrongly tagged and tagged events as well as the total number of events in the B^+ data are obtained from the simultaneous fit on the B^+

mass distribution and counted directly from weighted numbers of tag leptons for the simulated samples. The maximum-likelihood fits on the B^+ mass distribution for the muon sample are shown in Figure 6.3. By comparing the tagging results with the baseline performance in Table 6.2, the mistag fraction of simulated B_s^0 events is decreased from 45% to about 29% for the muons, and from 47% to about 32% for electrons, and the tagging powers are raised from 0.19% to 0.67% and 0.05% to 0.38%, respectively. Thus, the cuts set on the lepton variables significantly improve the tagging performance.

The wrong tag fractions, tagging efficiencies and tagging powers determined from the simulated B_s^0 and B^+ events agree within roughly two standard deviations of the uncertainties for both lepton taggers. The B^+ data has a slightly larger wrong tag fraction and better tagging efficiency compared to the simulations. These two differences balance each other leading to a similar tagging power compared to the simulated samples.

The deviations between the tagging efficiencies in the data and simulated samples can be due to the differences in the b hadron production processes in the simulations and real world. The data contain more $b\bar{b}$ events from the gluon splitting process and fewer events from the flavour excitation production mechanism than the simulated samples [86]. Tagging efficiency of the $b\bar{b}$ events produced in the flavour excitation process is smaller with respect to other $b\bar{b}$ events since the pseudorapidity of the opposite-side b hadron is more likely to decay outside the detector acceptance. As the number of gluon splitting events is underestimated and flavour excitations overestimated in the simulated samples, lower tagging efficiencies are obtained in the simulations compared to the data.

Another reason for the deviations in tagging performances between the data and simulated samples can be due to relatively old versions of PYTHIA and EvtGen generators used to simulate $b\bar{b}$ production and the b hadron decay processes. In

fact, the b hadron production and branching fractions are slightly different in the used PYTHIA and EvtGen generators compared to the latest measurements. This affects mainly the wrong tag fraction determined from the simulated samples.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B^+ \to J/\psi K^+$ data
wrong tag fraction ω [%]	29.0 ± 0.4	28.3 ± 0.3	30.7 ± 0.4
tagging efficiency ϵ_{tag} [%]	3.79 ± 0.03	3.87 ± 0.03	4.55 ± 0.03
tagging power P_{tag} [%]	0.67 ± 0.02	0.73 ± 0.02	0.68 ± 0.03

Table 6.4: Tagging performance of muons in the simulated and weighted B_s^0 and B^+ samples and the B^+ data. The uncertainties shown in the table are statistical.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B^+ \to J/\psi K^+$ data
wrong tag fraction ω [%]	31.6 ± 0.7	33.2 ± 0.4	34.8 ± 0.3
tagging efficiency ϵ_{tag} [%]	2.84 ± 0.04	2.81 ± 0.02	3.26 ± 0.02
tagging power P_{tag} [%]	0.38 ± 0.02	0.32 ± 0.02	0.30 ± 0.02

Table 6.5: Electron tagging performance of the simulated and weighted B_s^0 and B^+ events and the B^+ data. The uncertainties shown in the table are statistical.

Binned and parametrised mistag fraction

The event-by-event mistag information is used in the weak phase analysis to increase the sensitivity of the ϕ_s measurement (see Equations (2.109)-(2.110)). The per-event mistag fraction is determined by first binning and then parametrising the wrong tag fraction as a function of the tag lepton p_{T}. The binned mistag fractions from the data events are obtained by performing a maximum-likelihood fit to the categories of mistagged and tagged events in all the lepton p_{T} bins simultaneously. The binned wrong tag fractions are parametrised using the function

$$\omega(p_{\mathrm{T}}) = p_0 + p_1 \exp\left(\frac{1}{p_2 + p_{\mathrm{T}}}\right), \tag{6.6}$$

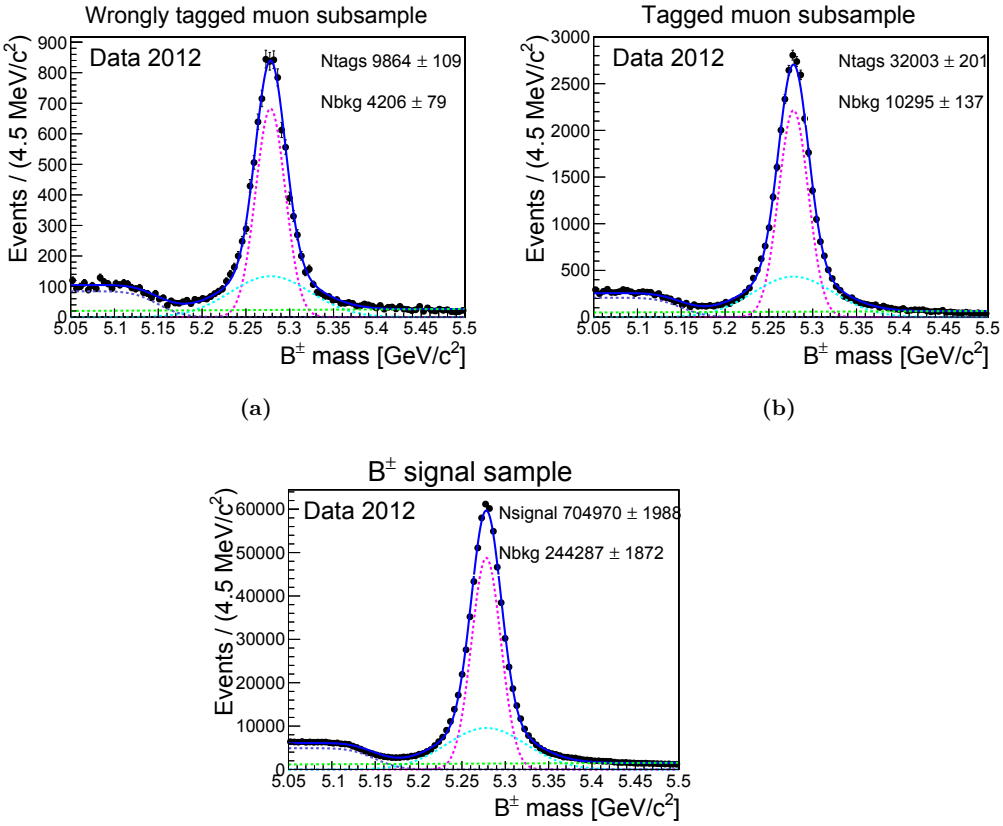


Figure 6.3: Numbers of a) wrongly tagged and b) tagged leptons as well as $B^+ \to J/\psi K^+$ signal events determined from 2012 data using the simultaneous fit on the B^+ mass distribution. Two Gaussian signal functions are plotted with dashed magenta and cyan lines. The background function formed of the error function and the first order Chevychev polynomial are described by dashed violet and green lines, respectively.

where p_T is the transverse momentum of the tag lepton. The coefficients p_i are fitted to the electron and muon mistag fraction distributions using a χ^2 technique. The values of the fitted parameters are shown in Tables 6.6 and 6.7. Binned and parametrised mistag fraction distributions are shown in Figure 6.4.

The binned average tagging power of the electron and muon taggers (see Equation (2.103)) are found to be $P_{tag} = 0.33 \pm 0.03\%$ and $P_{tag} = 0.72 \pm 0.02\%$, respectively. No significant improvement with respect to the average tagging performances is obtained from the binned approach.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B^+ \to J/\psi K^+$ data
p_0	-0.4 ± 0.2	-0.3 ± 0.1	-0.2 ± 0.1
p_1	0.6 ± 0.2	0.4 ± 0.1	0.4 ± 0.1
p_2	1.2 ± 0.7	-0.4 ± 0.5	-0.2 ± 0.3

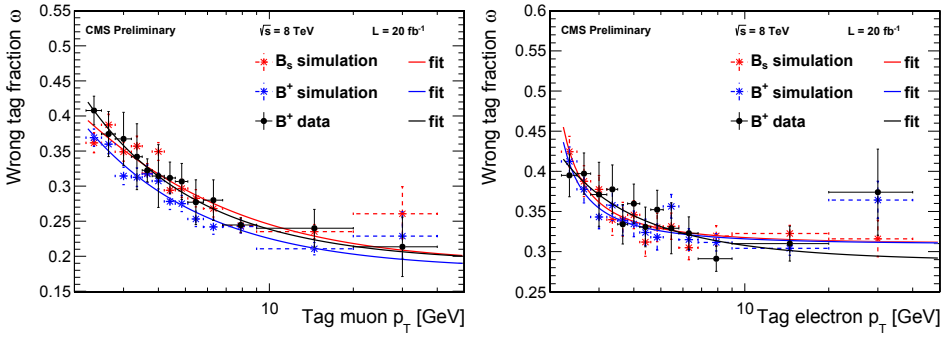
Table 6.6: Coefficients p_i for parametrised muon mistag fraction.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B^+ \to J/\psi K^+$ data
p_0	0.22 ± 0.04	0.22 ± 0.03	0.18 ± 0.04
p_1	0.08 ± 0.04	0.07 ± 0.03	0.11 ± 0.03
p_2	-1.1 ± 0.3	-1.2 ± 0.3	-0.8 ± 0.3

Table 6.7: Coefficients p_i for parametrised electron mistag fraction.

Combined average tagging performance

The two taggers are combined to determine the final average tagging performance. The lepton with a higher value of absolute dilution $|D| = |1 - 2\omega|$ is chosen to tag the event in those cases where the events contains both lepton tags. The per-event wrong tag fractions are obtained from the mistag fraction parametrisations of the both lepton species. No correlation between the tag leptons is considered as the two lepton species are mutually exclusive. Only 2.9% of the tagged events contain both lepton species. The combined tagging performance is determined by performing a simultaneous fit to the wrongly tagged, tagged and signal events of



(a) **(b)**

Figure 6.4: Binned mistag fractions for a) muons and b) electrons as a function of lepton p_{T}. The uncertainties are statistical only.

the $\mathrm{B}^+ \to \mathrm{J}/\psi\mathrm{K}^+$ mass distribution. The combined average tagging performance is $\omega = (32.2 \pm 0.3)\%$, $\epsilon_{tag} = (7.67 \pm 0.04)\%$ and $P_{tag} = (0.97 \pm 0.03)\%$, where the uncertainties are statistical only.

6.7 Neural network-based tagging algorithm

The neural network-based tagging algorithm is developed at University and INFN Padua and is fully documented in Reference [76]. Even if the algorithm itself is created in Padua, the tagging performance of the NN-tagger is validated with B^0_d data at University of Helsinki. As the validation procedure is thoroughly discussed in this thesis, also a short description of the NN algorithm itself and its tagging results is given here.

6.7.1 Multilayer perceptron network

A multilayer perceptron neural network (see e.g. [87] p. 156-166) is a set of connected neurons that each produce a response to the given vector of input variables. The network maps the n-dimensional input space to a lower dimensional space of

the output variables. The response of the network is determined by the weights of the connected neurons.

A network consists of an input and an output layers and an arbitrary number of hidden layers in between. Each layer further comprises a collection of neurons. Input layer has exactly the same number of neurons as there are input observables. The number of neurons in the output layer depends on the nature of the machine learning task. In the signal-background classification there is typically only one output variable, which describes the signal-likeness of a given data event. The number of hidden layers and number of neurons inside can vary and has to be optimised to obtain the best classification performance.

The neurons give a response to a weighted sum of input instances. The next layer of neurons takes the weighted sums of the responses of the previous layer as an input and the signal propagates throughout the network. A logistic function and a hyperbolic tangent are commonly used as response functions. An illustration of the MLP is shown in Figure 6.5.

A network learns the weights of the neurons using a back propagation algorithm [87]. First, random weights are cast to each neuron, and the output is computed for the training events. For each training instance, a mean squared error i.e. loss of the network is determined by comparing the output of the network to the correct output. The second step has a backward pass through the network. During the pass, the partial derivatives of the loss function with respect to the weights are determined and the weights are shifted towards the negative gradient of the loss function. The process is repeated until the performance of the algorithm no longer improves.

6.7.2 Lepton preselection

The muons and electrons used in the MLP tagger are reconstructed with the PF algorithm and a preselection is applied to the leptons before optimising the neural

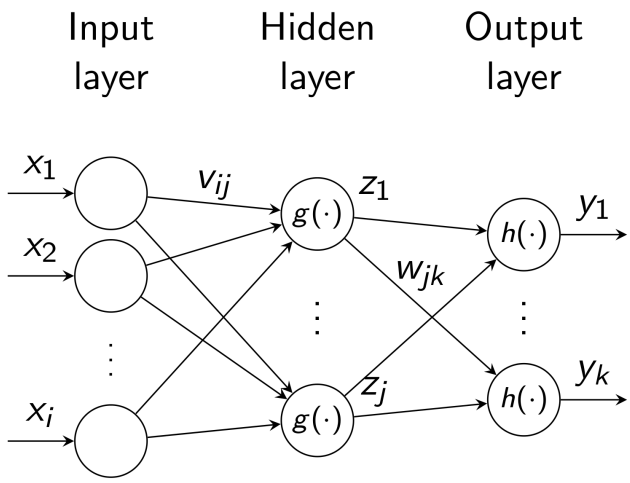


Figure 6.5: An illustration of a MLP network that has one hidden layer. x_i are the input features and V_{ij}, W_{jk} are the weights. Neurons z_j use a response function $g(\cdot)$. The output layer uses a different response function $h(\cdot)$, and the network has several output variables y_k. Figure is taken from the course material of a machine learning course at University of Helsinki [88].

networks. Selection is done in order to decrease the number of leptons that do not originate from the decays of opposite-side b hadrons. Several cut configurations are tried using the same observables deployed by the cut-based tagger, but also including lepton pseudorapidity and PF isolation variable in the variable set. The best tagging powers result from the cut configurations obtained from the optimisation of the cut-based tagging algorithm.

Thus, the preselection for muons is the same as shown in Table 6.3. The IP, p_{T}, and ΔR requirements for the tag electrons are also exactly the same as shown in Table 6.3. As the electron category is changed between the tagging algorithms, the electron identification discriminator is switched to a discriminator (PFmva) designed to decrease the fake rate of the PF electrons. A preselection cut on the PFmva discriminator is optimised to be greater than 0.2.

6.7.3 Input variables

Various input observables are given to the network in order to find the best classification success rate for the leptons originating from b $\rightarrow$ lX decays. The inputs used in the cut-based algorithm are also fed into the neural networks. However, the angular distance ΔR between the lepton and the b hadron is dropped from the final version of the network as its discriminating power was shown to be worse when compared to another set of observables. The list of the variables used in the final version of the network is given below.

- Lepton p_{T} and pseudorapity η

  As already described in Section 6.6.1, the kinematic observables of the leptons are helpful for rejecting the leptons and other particles coming from background processes. Misidentified leptons (e.g. kaons and pions) are more often produced in the forward region of the detector and thus have high values of pseudorapidity compared to the leptons from semileptonic b-hadron decay. In addition, the transverse momentum of misidentified particles can also be lower than the p_{T} of true tag leptons.

- Three-dimensional impact parameter with respect to the lepton and the b hadron

  The motivation to use this parameter is given already in Section 6.6.1.

- Momentum relative to the axis of the jet associated to the lepton ($p_{\mathrm{T}}^{\mathrm{rel}}$)

  The observable is defined as $p_{\mathrm{T}}^{\mathrm{rel}} = p_l \sin\theta$, where $\cos\theta = \frac{\mathbf{p}_l \cdot (\mathbf{p}_{jet} - \mathbf{p}_l)}{p_l |\mathbf{p}_{jet} - \mathbf{p}_l|}$. The quantity can only be determined in the presence of a jet associated with the lepton. The jets used to obtain the $p_{\mathrm{T}}^{\mathrm{rel}}$ observable belong to a loose PF jet category [89]. The selection requirements of the loose PF jets have been slightly reoptimised after the publication of the paper, but this is not of great importance for flavour tagging. A schematic illustration of the lepton $p_{\mathrm{T}}^{\mathrm{rel}}$ is

shown in Figure 6.6. A minimum angular distance of $\Delta R > 0.5$ between the B meson flight direction and the jet momentum is imposed to decrease the number of jets that contain the signal-side B meson.

The p_T^{rel} observable reflects the momentum of the muon in the B meson rest frame. The muons from b $\rightarrow$ lX decays often have higher values of p_T^{rel} due to the larger mass of b quarks compared to the leptons originating from the decays light quarks or b $\rightarrow$ cX $\rightarrow$ lX$'$ process.

- Lepton charge cone (Q_l)

  In addition to the lepton itself, the lepton charge cone is another observable used to determine the charge of the opposite-side b hadron. The charge cone is defined as

  $$Q_l(\Delta R, k, l-in/l-out) = \frac{\sum_{i=1}^{N_{tracks}} q^i (p_T^i)^k}{\sum_{i=1}^{N_{tracks}} (p_T^i)^k}, \tag{6.7}$$

  where the sum spans over all the charged PF candidates found within the cone ΔR around the lepton. Parameter q^i is the charge of the i-th PF candidate in the sum. The tracks of PF particles must have $p_T^i > 0.5$ GeV, $|\eta^i| < 2.5$ and a minimum of five hits in the tracker is required. The tracks originating from the signal side decay are excluded from the sum. The cone size ΔR, the exponent k and the inclusion/exclusion of the lepton itself in the cone ($l-in/l-out$) are tuned to obtain the best discriminating power between the tag leptons and leptons from background processes. Although both $l-in, l-out$ definitions are used in the muon NN, generally the definition with the lepton inside the cone leads to better discrimination of electrons and muons compared to $l-out$ option.

- PF isolation (PFiso)

  The PF isolation [70] describes the isolation of the lepton from the surrounding detector activity and is used in the background rejection. The value of the

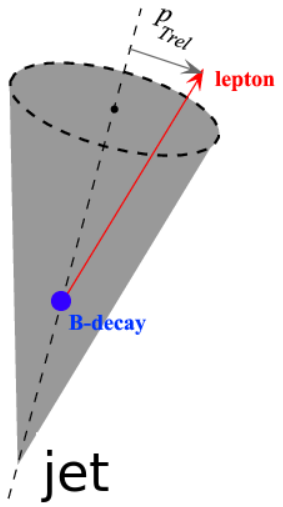


Figure 6.6: A schematic view of the lepton $p_{\mathrm{T}}^{\mathrm{rel}}$ observable. Figure is taken from [91].

quantity is determined using the missing transverse energy (E_T) of the charged hadrons that are associated with the same primary vertex as the lepton. The missing energies of photons and neutral hadrons are taken into account within $\Delta\mathrm{R} < 0.4$ cone around the lepton flight direction. The PF isolation is defined as

$$\mathrm{PFiso} = \frac{\sum E_T(\text{charged hadrons from the same PV}) + \sum E_T(\text{neutral hadrons}) + \sum E_T(\text{photons})}{p_{\mathrm{T}}^{l}}, \quad (6.8)$$

where p_{T}^{l} is the transverse momentum of the lepton. The leptons coming from the $\mathrm{b} \to \mathrm{lX}$ transition have lower values of PF isolation than the leptons arising either from the cascade decays of B hadrons or from the background processes.

- Multivariate electron identification discriminator (PFmva)

  A multivariate electron discriminator is used to reduce the background particles misidentified as electrons. The discriminator is tuned to separate the electrons from other particles, especially from pions faking the electrons. More details about the discriminator are found in Reference [90].

6.7.4 Network structure

The MLP networks are trained using half of the leptons from the simulated $B_s^0 \to J/\psi\phi$ sample. Furthermore, the leptons are divided into three categories regarding the correlation between the flavour of the signal-B meson and the charge of the lepton. These categories are correctly charged leptons that arise from the semileptonic decays of opposite-side B mesons, wrongly charged leptons coming from $b \to cX \to lX'$ process and randomly charged leptons from non-B hadron decays. The training set consists of 12000 opposite-side muon events and 10200 electron events. The other halves of the lepton events are used to test the performances of the networks.

The networks are trained and tested with several sets of the input variables and varying the network structure to obtain the input variables and network configuration with the best signal vs background separation. The final set of the NN input variables is selected to maximize the integral of the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve presents the background rejection versus signal efficiency for a given test sample. The optimised network structures and training characteristics are listed in Table 6.8 and the input variables are shown in Table 6.9.

6.7.5 Tagging performance of the MLP algorithm

Average tagging performance

The average tagging performances of the lepton taggers are determined by setting a single cut on the MLP discriminators, $MLP_\mu > 0.64$ and $MLP_e > 0.64$. The resulting tagging performances are shown in Tables 6.10 and 6.11.

	Muon MLP	Electron MLP
Training cycles	750	600
Hidden layers	16	11
Response function	tanh	tanh
Neuron input type	sum	sum
Learning rate	0.02	0.02
Decay rate	0.01	0.01
Test rate	5	5

Table 6.8: The structure of the muon and electron networks.

Binned and parametrised tagging performance

The mistag fraction is also binned in twenty bins of the MLP discriminator. The binned tagging performances evaluated with the B^+ data are found to be $P_{tag} = 0.84 \pm 0.03\%$ and $P_{tag} = 0.48 \pm 0.02\%$ for muons and electrons, respectively. Lastly, the wrong tag fractions are parametrised as a function of the MLP outputs. A modified error function defined as

$$\omega(\text{MLP}) = p_0 + p_1 \cdot [1 - \text{Erf}(p_2 + p_3 \cdot \text{MLP})] \tag{6.9}$$

is used in the parametrisation. The parametrised mistag fractions for electron and muon taggers are shown in Figure 6.7.

The parametrised mistag fractions ($\omega^{calc} = \omega(\text{MLP})$) are finally calibrated using the measured mistag fractions (ω^{meas}) in the bins of MLP output. In the absence of biases affecting the mistag parametrisation, a linear dependence between the measured and the calculated ω should be observed. The calibration line is defined as

$$\omega^{meas}(\omega^{calc}) = p_0 + p_1 \cdot (\omega^{calc} - \omega'), \tag{6.10}$$

where ω' is a constant set to 0.35. The above parametrisation for the calibration

	Muon MLP	Electron MLP
Lepton p_{T}	x	x
Lepton η	x	x
Lepton 3D IP	x	x
Lepton PF isolation	x	x
$Q_l(\Delta\mathrm{R}, k, l-in)$	x	x
$Q_l(\Delta\mathrm{R}, k, l-out)$	x	-
Lepton $p_{\mathrm{T}}^{\mathrm{rel}}$	x	-
PF electron discriminator	-	x

Table 6.9: The input variables used in the muon and electron networks. The charge cones variables $Q_\mu(\Delta\mathrm{R}=0.5, k=1.50, \mu-in)$, $Q_\mu(\Delta\mathrm{R}=0.5, k=1.10, \mu-out)$ and $Q_e(\Delta\mathrm{R}=0.3, k=1.75, e-in)$ are exploited for muons and electrons, respectively.

is used instead of the usual first-level polynomial in order to reduce the correlation between the coefficients p_0 and p_1. The calibration functions obtained for the muon and electron taggers are shown in Figure 6.8, and the resulting calibration parameters are listed in Table 6.12. Although the fitted parameters are consistent with the hypothesis of linear dependence and zero bias, the electron and muon mistag fractions obtained from parametrisations are corrected using the calibration curves. The tagging performances with parametrised mistag fractions are presented in Table 6.13 where the first uncertainty is statistical and the second one is systematic.

Combined tagging performance is determined in a similar fashion as for the individual electron and muon taggers. The overlap between the lepton taggers is removed by tagging the event with the lepton for which the absolute value of per-event dilution is the highest. The combined tagging performance is also shown in Table 6.13.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B_d^0 \to J/\psi K^{0*}$ MC	$B^+ \to J/\psi K^+$ Data
ϵ_{tag} [%]	2.81 ± 0.02	2.86 ± 0.02	2.92 ± 0.03	3.34 ± 0.02
ω [%]	23.1 ± 0.3	23.5 ± 0.3	23.5 ± 0.4	25.9 ± 0.4
P_{tag} [%]	0.82 ± 0.02	0.80 ± 0.02	0.80 ± 0.02	0.78 ± 0.03

Table 6.10: Muon tagging performances evaluated on the $B_s^0 \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ simulations, and on the $B^+ \to J/\psi K^+$ channel of the 2012 data. A cut on the MLP discriminator $MLP_\mu > 0.64$ is applied.

	$B_s^0 \to J/\psi\phi$ MC	$B^+ \to J/\psi K^+$ MC	$B_d^0 \to J/\psi K^{0*}$ MC	$B^+ \to J/\psi K^+$ Data
ϵ_{tag} [%]	1.90 ± 0.02	1.91 ± 0.02	1.96 ± 0.02	2.11 ± 0.02
ω [%]	25.5 ± 0.4	26.4 ± 0.5	25.5 ± 0.5	27.5 ± 0.5
P_{tag} [%]	0.45 ± 0.02	0.42 ± 0.02	0.47 ± 0.02	0.43 ± 0.02

Table 6.11: Electron tagging performances evaluated on the $B_s^0 \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ simulations, and on the $B^+ \to J/\psi K^+$ channel of the 2012 data. A cut on the MLP discriminator $MLP_e > 0.64$ is applied.

Systematic uncertainties

Several sources of systematic uncertainties are considered to affect the MLP tagging performance, and their contributions to total systematic uncertainty are quantified.

- Fit model for the B^+ mass fit

  The binned mistag fraction of $B^+ \to J/\psi K^+$ data is determined by evaluating the numbers of mistagged and tagged events using a likelihood fit to the B^+ mass distribution. The effect of the fit model on the wrong tag fraction and tagging efficiency is quantified. First the signal pdf is changed from a double-Gaussian function to a sum of three Gaussians, and the residual difference of the mistag fractions and tagging efficiencies between the nominal and the alternated fit is evaluated. Then the first-order Chebyshev polynomial in the

	Muon tagger	Electron tagger
p_0	0.350 ± 0.004	0.350 ± 0.004
p_1	1.00 ± 0.04	1.01 ± 0.04

Table 6.12: Calibration coefficients of the muon and electron taggers. The uncertainties are statistical only.

	Muons	Electrons	Combined tagger
ϵ_{tag} [%]	4.56 ± 0.02	3.92 ± 0.02	8.31 ± 0.03
ω [%]	$28.64 \pm 0.32 \pm 0.015$	$32.45 \pm 0.36 \pm 0.05$	$30.2 \pm 0.2 \pm 0.05$
P_{tag} [%]	$0.833 \pm 0.024 \pm 0.012$	$0.48 \pm 0.020 \pm 0.003$	$1.307 \pm 0.031 \pm 0.007$

Table 6.13: Parametrised tagging performances evaluated on the $B^+ \to J/\psi K^+$ sample of the 2012 data for electrons, muons, and for the combined single-lepton tagger.

background model is switched to an exponential function and the difference with respect to nominal results is quantified. Even if the yields of mistagged and tagged events change between the fits, the values of binned mistag fraction and the tagging efficiency are consistent within the nominal and alternated fit models. Hence, no systematic uncertainty is assigned due to the B^+ fit model.

- Data taking period

  The effect of different pileup conditions on tagging performance is evaluated. The data is divided into three subsamples of data-taking periods. The wrong tag fractions and the mistag calibration parameters for each period are determined. As the calibration parameters and hence also the mistag fractions are consistent in each data-taking period, a systematic uncertainty is not assigned.

- B^+ production properties

  The $B^+ \to J/\psi K^+$ events are divided into three subsamples according to the

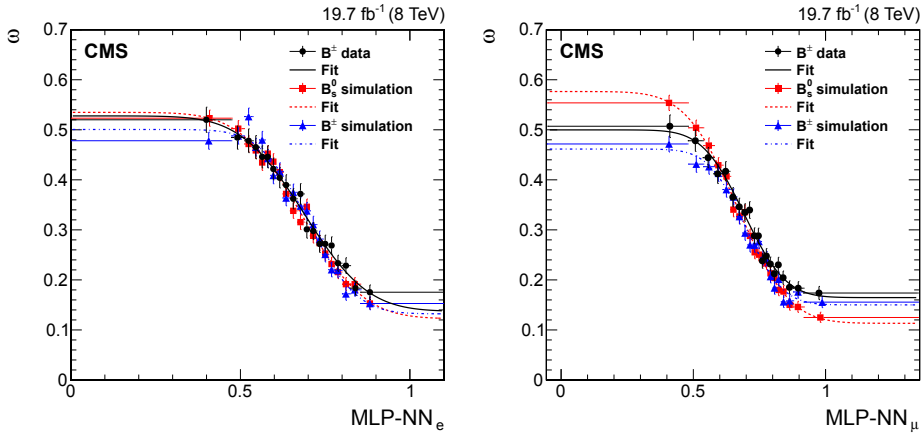


Figure 6.7: Mistag probability ω for electrons (left) and muons (right) as a function of multiLayer perceptron neural network output (MLP-NN). Figure is taken from [76].

transverse momentum of the reconstructed B^+ meson to quantify possible dependence between the mistag fraction and the production properties of the signal b hadron. The p_T regions used to split the data are $p_T < 18$ GeV/c, $18 \leq p_T < 26$ GeV/c and $p_T > 26$ GeV/c. The mistag calibration parameters for each subsample are determined and found to be statistically in agreement. Thus, no systematic uncertainty is accounted for the B^+ production properties.

- Signal-side b hadron species

  The uncertainty related to the signal-side b hadron species is estimated in the following manner: The mistag fraction parametrisation (ω^{calc}) is taken from the $B^+ \to J/\psi K^+$ sample and the measured wrong tag fractions ω^{meas} are taken from the $B_s^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{0*}$ MC samples. Two sets of $\omega^{calc}(\omega^{meas})$ calibration data are obtained in bins of measured mistag fraction, and the data are fitted with the calibration function. The new calibration curves are used to correct for the parametrised ω^{calc} obtained from the

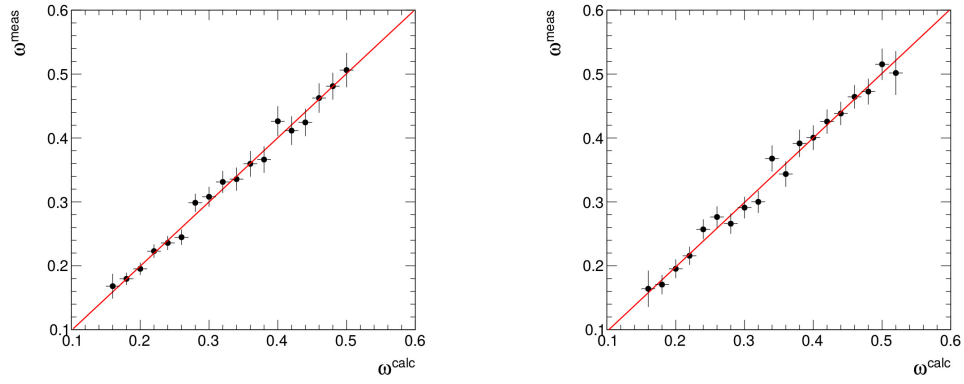


Figure 6.8: The measured mistag fraction as a function of the parametrised wrong tag fraction for muons (left) and electrons (right). Figure is taken from [76].

$B^+ \to J/\psi K^+$ data. New mistag fractions and tagging powers are determined. The maximum difference of the new mistag fractions and tagging powers with respect to the nominal ones is taken as a systematic uncertainty.

The uncertainties related to the flavour dependence of muon and electron tagging performance are found to be $\sigma_{syst,\omega} = 0.15\%$ and $\sigma_{syst,P_{tag}} = 0.012\%$ and $\sigma_{syst,\omega} = 0.05\%$ and $\sigma_{syst,P_{tag}} = 0.003\%$, respectively. The uncertainties for the combined tagging performance are $\sigma_{syst,\omega} = 0.05\%$ and $\sigma_{syst,P_{tag}} = 0.007\%$.

- Mistag fraction parametrisation

  The binned wrong tag fraction distributions are parametrised with two additional functions

$$\begin{aligned}\omega(\text{MLP}) &= p_0 + p_1 \cdot [1 - \arctan(p_2 + p_3 \cdot \text{MLP})] \\ \omega(\text{MLP}) &= p_0 + p_1 \cdot [1 - \tanh(p_2 + p_3 \cdot \text{MLP})].\end{aligned}$$

  The average mistag fractions and the tagging performances are determined using the new parametrisations. The maximum residual difference of 10^{-4} between the mistag fractions and tagging powers is negligible compared to the

uncertainty originating from the signal-side b hadron species.

6.8 B_d^0 mixing asymmetry

In addition to the $B^+ \to J/\psi K^+$ data, the performance of MLP algorithm can be validated with the reconstructed $B_d^0 \to J/\psi K^{0*}$ events. It is important to confirm that the tagging performance and especially the mistag fraction is not significantly different among the b hadron species. If that should be the case, the tagging results obtained with B^+ mesons would not hold for B_s^0 events.

A value of the mass difference in the B_d^0 eigenstates can be obtained as a side product of the tagging validation. The flavour oscillations are revealed by raw mixing asymmetry that shows how yields of mixed and unmixed B_d^0 ($\bar{B}_d^0$) mesons vary as a function of the measured proper decay time. Experimentally the raw mixing asymmetry is defined as

$$A_{raw}(t) = \frac{N_{unmixed}(t) - N_{mixed}(t)}{N_{unmixed}(t) + N_{mixed}(t)}, \tag{6.11}$$

where $N_{unmixed}(t)$ and $N_{mixed}(t)$ are yields of mixed and unmixed events at time t. The raw asymmetry is proportional to the mixing asymmetry (see Section 2.4.3)

$$A_{mix}(t) \approx (1 - 2\omega_{fit}) \cos(\Delta m_d t), \tag{6.12}$$

where Δm_d is the mass difference of the B_d^0 eigenstates and ω_{fit} is the mistag fraction associated to the B_d^0 flavour tagging.

The raw mixing asymmetry distribution is reconstructed using an SPlot technique [92]. It is possible to obtain rough estimates for the mass difference Δm_d and the mistag fraction ω_{fit} directly by fitting Equation (6.12) to the raw mixing asymmetry distribution. An improved precision for ω_{fit} and Δm_d can be obtained from a more complex fit to the tagged B_d^0 events. This strategy is exploited for measuring the mistag fraction of the B_d^0 events in this thesis.

6.8.1 Likelihood function for B_d^0 mixing asymmetry

The mistag fraction and the mass width difference are obtained by performing a three-dimensional unbinned extended maximum-likelihood fit to distributions of reconstructed B_d^0 mass, proper decay time and its uncertainty. The total likelihood function is

$$L_{tot} = N_{sig} L_{signal}(t, q, m, \sigma_t) + N_{bkg} L_{background}(t, q, m, \sigma_t), \tag{6.13}$$

where $L_{signal}(t, q, m)$ are the signal and background likelihoods respectively and N_{sig} and N_{bkg} are the numbers of signal and background events. The signal likelihood is defined as

$$L_{signal}(t, q, m, \sigma_t) = \prod_{i=1}^{N} P_{sig}(m_i) P_{Sig}(t_i) P_{sig}(\sigma_{t,i}) P_{sig}(q_i) P_{sig}(q_i') \tag{6.14}$$

where

- $P_{sig}(m) = \sum_{j=1}^{3} f_j \frac{1}{\sqrt{2\pi}\sigma_{m,j}} e^{-(m-\mu_m)^2/2\sigma_{m,j}}$ is a triple Gaussian function that describes the B_d^0 mass peak. The standard deviations of the Gaussians are denoted as $\sigma_{m,j}$ and the average B_d^0 mass is μ_m. The relative strength between the Gaussians are denoted as f_j and the strengths add up to unity: $\sum_{j=1}^{3} f_j = 1$.

- $P_{sig}(\sigma_t) = \frac{\sigma_t^{\gamma-1} e^{-\sigma_t/\beta}}{\Gamma(\gamma)\beta^\gamma}$ is a gamma function that models the decay time uncertainty distribution. The shape and scale parameters of the gamma function are denoted as β and γ.

- $P_{sig}(t) = [\frac{1}{2\tau_d} e^{-t/\tau_d}(1 + q(1 - 2\omega_{fit}) \cos(\Delta m_d t))] \otimes R(t)$ is an exponential decay time function times the mixing asymmetry convoluted with a Gaussian decay time resolution function $R(t)$. Parameter τ_d is the average lifetime of the B_d^0 meson. The resolution function is defined as $R(t) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-t^2/2\sigma_t}$, where the

width σ_t is taken from the proper decay time uncertainty distribution in an event-by-event fashion.

- $P_{sig}(q_i)$ and $P_{sig}(q'_i)$ are the mixing and tagging asymmetry functions, respectively. Argument q_i of $P_{sig}(q_i)$ determines the tag decision for mixed ($q_i = -1$) and unmixed ($q_i = 1$) state. In a similar fashion, argument q'_i of $P_{sig}(q'_i)$ determines the tag decision for B^0_d ($q'_i = 1$) and $\bar{B}^0_d$ ($q'_i = -1$) meson. Both functions contain only one free parameter that is the fraction between the mixed (B^0_d) and unmixed ($\bar{B}^0_d$) events, respectively.

The background component is defined as

$$L_{background}(t, q, m) = \prod_{i=1}^{N} P_{bkg}(m_i) P_{bkg}(t_i) P_{bkg}(\sigma_{t,i}) P_{bkg}(q_i) P_{bkg}(q'_i) \tag{6.15}$$

where

- $P_{bkg}(t) = [\sum_{k=1}^{2} f_k e^{-t/\tau_{bkg,k}}] \otimes R_{bkg}(t)$ is a sum of two exponential functions convoluted with a Gaussian resolution $R_{bkg}(t)$. Parameters $\tau_{bkg,k}$ are the lifetimes associated with the background decay time distribution. Parameter f_k is a relative strength between the exponentials and the strengths add up to unity, $\sum_{k=1}^{2} f_k = 1$. The resolution function is defined in the same manner as in the signal pdf.

- $P_{bkg}(\sigma_t)$ is a gamma function that models the background component of the decay time uncertainty distribution.

- $P_{bkg}(m) = e^{-Cm}$ is an exponential function that describes the background component of the mass distribution. Coefficient C is a slope of the exponential.

- $P_{bkg}(q_i)$ and $P_{bkg}(q'_i)$ are the mixing and tagging asymmetry functions for the background component.

6.8.2 Cross-checks with simulated B_d^0 sample

In order to ensure that the MLP tagger is properly integrated in the raw mixing asymmetry analysis, the tagging performances obtained from the simulated $B_d^0 \to J/\psi K^{0^*}$ sample are compared to the results reported for the NN tagger. The tagging performances obtained in the framework of mixing asymmetry study are shown in Table 6.14. The tagging results reported for the MLP algorithm presented in Tables 6.10 and 6.11 are in agreement with the performances shown in Table 6.14.

[%]	Muons	Electrons
ω	24.1 ± 0.2	25.1 ± 0.3
ϵ_{tag}	2.93 ± 0.01	2.01 ± 0.01
P_{tag}	0.79 ± 0.01	0.50 ± 0.01

Table 6.14: Muon and electron tagging performances evaluated for the simulated $B_d^0 \to J/\psi K^{0^*}$ events by cutting on the electron and muon MLP discriminators, $MLP_e > 0.64$ and $MLP_\mu > 0.64$.

6.8.3 Mistag fraction in $B_d^0 \to J/\psi K^{0^*}$ decay

The validation of the tagging performance using $B_d^0 \to J/\psi K^{0^*}$ events requires identifying mixed and unmixed B_d^0 events. It is worth noting that both the production time and decay time flavours of the B_d^0 meson have to be tagged in order to declare the event mixed or unmixed.

The flavour at the production time is given by the tagging algorithm while the decay time flavour is obtained from the mass hypothesis of the $K^*(892)^0$ ($\overline{K}^*(892)^0$) candidate (see Section 6.3.4). Since $K^*(892)^0$ meson has a large decay width of 50 MeV, the hypothesis that gives a value closer to the $K^*(892)^0$ world-average mass is not always correct. Mistagged decay time flavour is found in 12.8% of the

simulated events. In order to compare the wrong tag fraction measured from the $B^+ \to J/\psi K^+$ decays and the fraction obtained from the mixing asymmetry fit, a dependence between the fitted wrong fraction and the final flavour mistags has to be quantified.

There are four different probabilities to tag the event:

- Probability to correctly (c) tag both the production and decay time flavours P(c $\to$ c)
- Probability to correctly tag the production time flavour and mistag (w) the decay time flavour P(c $\to$ w)
- Probability to mistag the production time flavour (w) and correctly tag the decay time flavour P(w $\to$ c)
- Probability to mistag both the production and decay time flavours P(w $\to$ w)

Naturally the four probabilities add up to unity:

$$\mathrm{P(c \to c) + P(w \to c) + P(c \to w) + P(w \to w) = 1.} \tag{6.16}$$

The fitted wrong tag fraction is the probability to have a mixed B_d^0 event in case of an unmixed event or vice versa. Since the measurement only uses information about the mixing, and the true B_d^0 flavours at the production and decay times are not considered, probability P(w $\to$ w) results in correctly tagged events in terms of mixing information. The fitted wrong tag fraction ω_{fit} is then a sum of two probabilities P(w $\to$ c) and P(c $\to$ w),

$$\omega_{fit} = \mathrm{P(w \to c) + P(c \to w) = 1 - P(c \to c) - P(w \to w).} \tag{6.17}$$

The probabilities P(c $\to$ c) and P(w $\to$ w) can be expressed using the initial flavour mistag fraction ω and the final flavour mistag fraction ω_f:

$$\mathrm{P(w \to w)} = \omega\omega_f \tag{6.18}$$

$$\mathrm{P(c \to c)} = (1-\omega)(1-\omega_f) \tag{6.19}$$

When the probabilities in Equations (6.18) and (6.19) are substituted in Equation (6.17), the fitted mistag fraction becomes

$$\omega_{fit} = \omega + \omega_f - 2\omega\omega_f. \tag{6.20}$$

6.8.4 Testing the fitted mistag fraction in $B_d^0 \to J/\psi K^{0^*}$ decay

A set of tests can be conducted with the simulated B_d^0 sample to ensure that the fitted mistag fraction behaves as derived in Equation (6.20), and that the likelihood fit works reliably. The first task is to determine the initial and flavour mistag fractions in two different ways and compare the results.

Both mistag fractions can be obtained by directly counting the simulated events with correctly and wrongly tagged events. On the other, hand the fractions can also be determined by fitting the likelihood function of Section 6.8.1 to two subsets of the simulated B_d^0 events. The final flavour mistag fraction can be obtained from a fit to events having correctly tagged initial flavour. In this case, the mistag fraction given by the fit corresponds to ω_f, as the initial flavour is always tagged correctly ($\omega = 0$). Similarly, the events having correctly tagged decay time flavour ($\omega_f = 0$) can be used to cross-check that the fitted mistag fraction coincides with production time mistag fraction since now $\omega_{fit} = \omega$. Finally, the fitted mistag fraction can be determined from Equation (6.20) when ω and ω_f are known. The result can then be compared with the result given by the likelihood fit to the full simulated sample.

The cross-checks are done using electrons and muons. No requirements are set for the values of the MLP discriminator in order to maximise the statistics of the lepton sample. It is worth noting that ω obtained from the fit to the simulated sample corresponds to the *unweighted* mistag fraction of combined electron and muon sample, and is not comparable to the reweighted wrong tag fractions presented in Section 6.8.2. The fit projections onto the mixing asymmetry are shown in

Figure 6.9. The wrong tag fractions obtained directly from simulation by counting mistagged events, fitting the likelihood function and predicted with Equation (6.20), are shown in Table 6.15. Mistag fractions determined with both methods and the corresponding prediction are consistent. The results show that the dependencies between fitted wrong tag fraction and the production and decay time mistag fractions are well understood. The fit also reproduces the mistag fraction obtained directly counting the correctly and wrongly tagged events with high accuracy.

[%]	counted from simulation	fitted over simulated events
$\omega_{unweighted}$	30.7 ± 0.2	30.7 ± 0.4
ω_f	12.81 ± 0.03	13.0 ± 0.3
expected ω_{fit} from Equation (6.20)	35.6 ± 0.1	35.7 ± 0.3
ω_{fit} obtained from counting/fitting	35.6 ± 0.2	35.3 ± 0.4

Table 6.15: Mistag fractions obtained by counting the simulated events, given by the fit and predicted by Equation (6.20). The uncertainties are statistical only.

6.8.5 Other cross-checks with the simulated B_d^0 sample

In order to further validate the fit procedure and an unbiased signal selection, the physics parameters τ_d and Δm_d obtained from the fit are compared to their corresponding generator-level values. The generator-level parameters and the fit results are shown in Table 6.16. The fit results are found to be consistent with the generator-level values, which further verifies the validity of the fit procedure and the signal selection.

6.8.6 Tagging results with $B_d^0 \to J/\psi K^{0^*}$ data

The mistag fractions are determined from the fits performed to the electron and muon samples separately. In addition, the total number of B_d^0 signal events is ex-

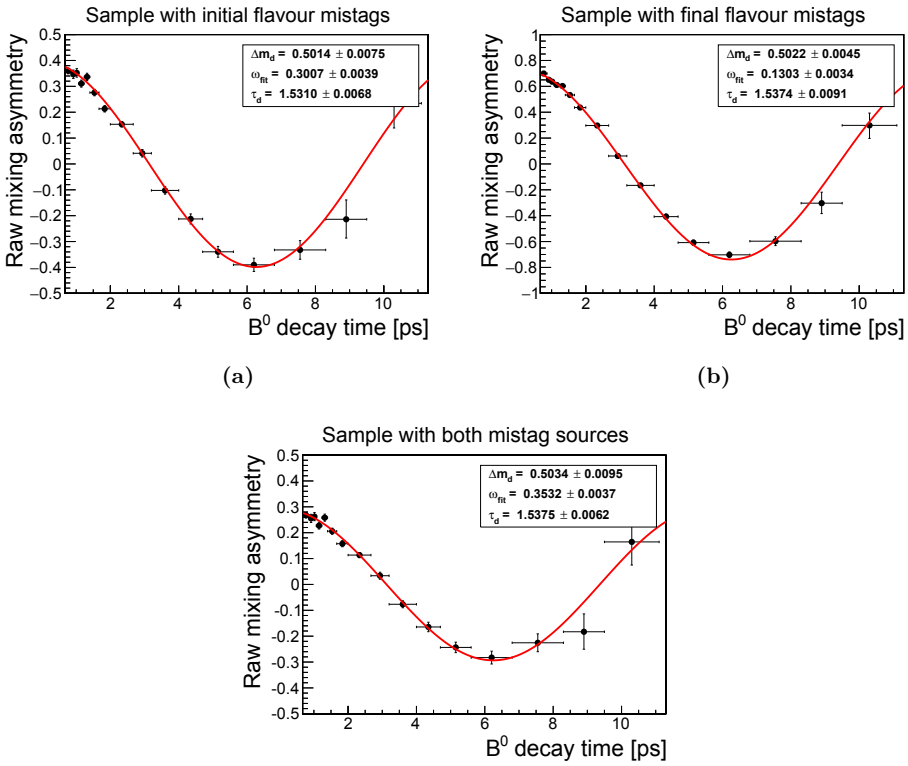


Figure 6.9: Asymmetry fit tested with three categories of simulated B_d^0 events: a) events with initial flavour mistags only, b) events with final flavour mistags only, and c) all the tagged B_d^0 events.

	MC gen. level value	Fit result
τ_d [ps]	1.536	1.538 ± 0.006
Δm_d [ps^{-1}]	0.502	0.503 ± 0.010

Table 6.16: The generator-level and fitted average lifetime and mass difference of the simulated $B_d^0 \to J/\psi K^{0^*}$ sample.

tracted from the B_d^0 mass fit shown in Figure 6.10. The yields of tagged events are shown in Table 6.17. The tagging performances of electrons and muons are reported in Tables 6.18 and 6.19 and compared to the tagging results obtained from the $B^+ \to J/\psi K^+$ sample. The electron tagging performances obtained from reconstructed B_d^0 and B^+ events are found to be consistent. A small discrepancy is observed in the muon tagging performance. The discrepancy may result from a statistical fluctuation the fit into muon sample, since the muon tagging results agree within 1.5σ uncertainties.

	$B_d^0 \to J/\psi K^{0^*}$ data
Number of events with electron tag	6091 ± 83
Number of events with muon tag	9850 ± 108
Total number of events	290143 ± 1366

Table 6.17: Yields total number of B_d^0 events and tagged events measured from the 2012 data.

	OS-μ tagger	
[%]	$B_d^0 \to J/\psi K^{0^*}$ data	$B^+ \to J/\psi K^+$ data
ω	23.7 ± 1.4	25.9 ± 0.4
ϵ_{tag}	3.39 ± 0.04	3.34 ± 0.02
P_{tag}	0.94 ± 0.10	0.78 ± 0.03

Table 6.18: Muon tagging performances evaluated for $B_d^0 \to J/\psi K^{0^*}$ and $B^+ \to J/\psi K^+$ samples of the 2012 data with a cut on muon MLP discriminator, $MLP_\mu > 0.64$.

6.8.7 Physics results from the asymmetry fit

The mass difference and the average lifetime of B_d^0 eigenstates are obtained from a combined lepton sample by fitting the likelihood of Section 6.8.1 to tagged B_d^0 events

	OS-e tagger	
[%]	$B^0_d \to J/\psi K^{0*}$ data	$B^+ \to J/\psi K^+$ data
ω	27.3 ± 1.8	27.5 ± 0.5
ϵ_{tag}	2.10 ± 0.03	2.11 ± 0.02
P_{tag}	0.43 ± 0.07	0.43 ± 0.02

Table 6.19: Electron tagging performances evaluated for the $B^0_d \to J/\psi K^{0*}$ and $B^+ \to J/\psi K^+$ samples of the 2012 data with a cut on electron MLP discriminator, $MLP_e > 0.64$.

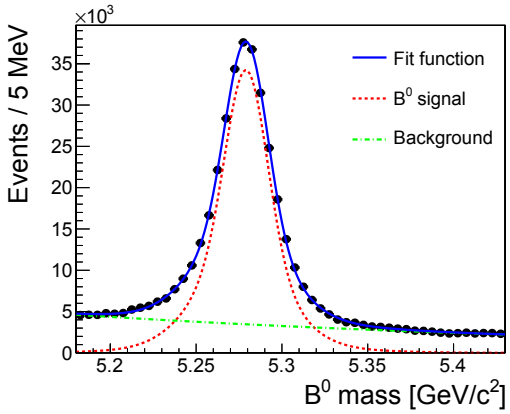


Figure 6.10: A fit to the total number of reconstructed B^0_d events.

using the cuts on MLP discriminators, $MLP_e > 0.64$ and $MLP_\mu > 0.64$. The fit projections of the B^0_d mass, decay time and decay time uncertainty distributions and the resulting mixing asymmetry are shown in Figure 6.11. The mixing asymmetry distribution is reconstructed from the weighted signal sample obtained with the SPlot technique. The signal and background weights are determined from the fit to the B^0_d mass distribution.

The result $\Delta m_d = 0.507 \pm 0.016$ ps^{-1} is in agreement with the previous measurements and the world-average value 0.5065 ± 0.0019 ps^{-1} [2]. The average lifetime is measured to be $\tau_d = 1.549 \pm 0.015$ ps^{-1}. The result is within two standard devi-

ations of the world-average value 1.520 ± 0.004 ps^{-1} [2].

The B_d^0 event reconstruction and the asymmetry fit are developed solely for the flavour tagging purposes, and are not optimised for the lifetime measurement. The average lifetime of the B_d^0 eigenstates can be measured more precisely using the full dataset of untagged B_d^0 events, as done in Reference [93].

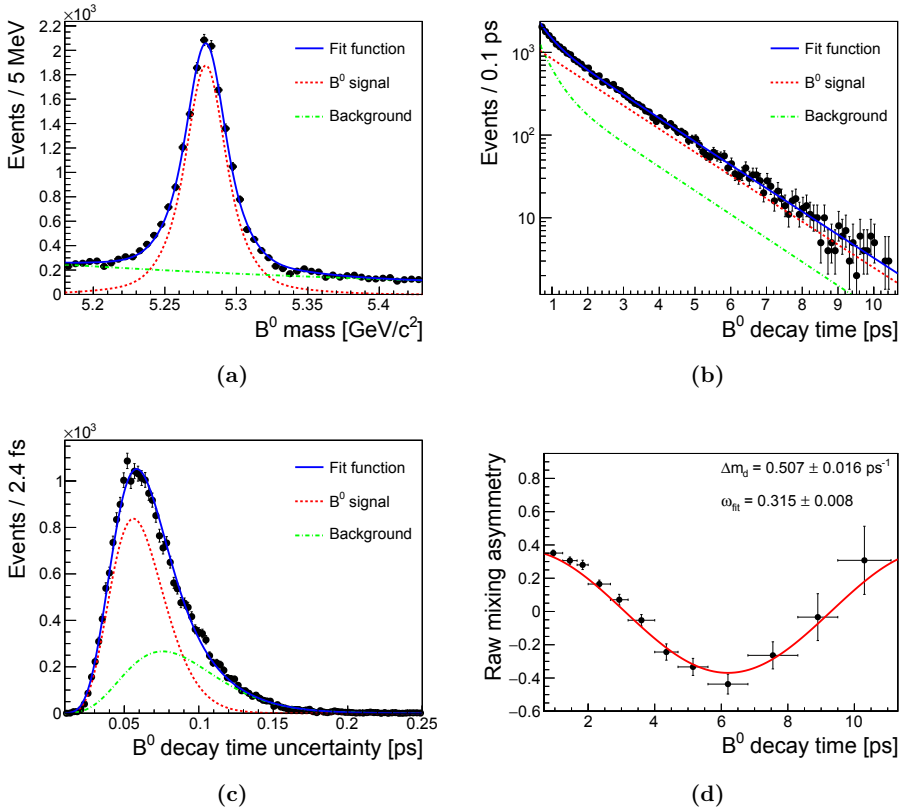


Figure 6.11: The projections of the final three-dimensional asymmetry fit to a) mass b) decay time and c) decay time uncertainty distributions. The solid blue line represents the total fit function, dotted red line is the signal component and green dashed-dotted line is the background component of the fit. The signal component of raw mixing asymmetry is shown in subfigure d).

6.9 Summary and outlook of the tagging studies

This chapter summarised the flavour tagging studies made in the CMS experiment. First, a simple cut-based tagging algorithm was developed, and the average and binned tagging performances with electron and muon tags were measured. Then the average tagging performance was measured by combining the leptons in a single tagging category. The first tagging algorithm was followed by a more sophisticated MLP algorithm that used the state-of-the-art machine learning techniques. The average, binned and parametrised tagging performances of the electron and muon taggers were determined. Moreover, the combined parametrised tagging performance was quantified.

The final tagging performance was determined using the $B^+ \to J/\psi K^+$ data. In addition, the tagging results of the NN tagger were validated with a measurement of the B_d^0 mixing asymmetry using $B_d^0 \to J/\psi K^{0^*}$ events reconstructed from the simulation and data. The data were collected by the CMS experiment in proton-proton collisions at a centre-of-mass energy of 8 TeV. The data correspond to an integrated luminosity of 19.7 fb^{-1}.

The tagging results obtained with the MLP algorithm outperform the results of the cut-based tagger. The binned muon tagging power ($P_{tag} = 0.84 \pm 0.03\%$) achieved with the NN tagger improves by 17% compared to the simple tagging algorithm ($P_{tag} = 0.72 \pm 0.02\%$). An improvement of 45% in the binned electron tagging power ($P_{tag} = 0.48 \pm 0.02\%$) is obtained compared to the simple tagger ($P_{tag} = 0.33 \pm 0.03\%$).

The flavour tagging studies were completed by developing a fitting procedure to measure the tagging performance from $B_d^0 \to J/\psi K^{0^*}$ data. This was done to cross-check the independence of the tagging results on the signal b-hadron species. The muon tagging performance obtained from the B_d^0 sample was found to be within 1.5 standard deviations of the results obtained from the B^+ sample. The electron

tagging performances of B_d^0 and B^+ events were found to be consistent. No dependency between the tagging performance and the b-hadron species was observed.

Furthermore, the mass difference Δm_d of the B_d^0 eigenstates is a free parameter of the asymmetry fit that was used to validate the tagging performance of the MLP algorithm. The mass difference given by the fit was $\Delta m_d = 0.507 \pm 0.016$ ps^{-1}, where the uncertainty is statistical only. The result is consistent with the world-average value 0.5065 ± 0.0019 ps^{-1} [2].

The total combined tagging power of the MLP lepton tagger is $P_{tag} = 1.307 \pm 0.031 \pm 0.007\%$ with the tagging efficiency of $\epsilon_{tag} = 8.31 \pm 0.03\%$ and the wrong tag fraction of $\omega = 30.2 \pm 0.2 \pm 0.05\%$, where the first uncertainty is statistical and the second is systematic. The LHCb experiment reports a total combined OS tagging power of $P_{tag} = 2.10 \pm 0.08 \pm 0.24\%$ for the $B^+ \to J/\psi K^+$ mode [42]. In a later weak mixing phase paper [94] they document an improved tagging power of $P_{tag} = 2.55 \pm 0.14\%$. The larger tagging power of the LHCb experiment in [42] compared to the CMS one originates from a vertex charge and opposite-side kaon taggers not present in the CMS tagging algorithm. The individual lepton tagging performances are similar in both experiments. The ATLAS experiment reports a total binned tagging power of $P_{tag} = 1.49 \pm 0.02\%$ [95] that includes jet charge and OS electron and muon taggers. The better total tagging performance compared to the CMS results is due to the jet charge tagger. The lepton taggers give similar performances compared to the CMS ones.

The flavour tagging algorithms have to be reoptimised for the forthcoming ϕ_s measurement that uses the data recorded with a center-of-mass energy of 13 TeV. Together with reoptimisation of the lepton taggers, the total tagging power can be improved by implementing the jet charge tagger. Since the CMS detector cannot distinguish between kaons and pions, implementing the same- or opposite-side kaon tagging will not largely increase the total tagging power.

Improvements in the tagging power of the opposite-side taggers may be further increased by including different input variables in the MLP network. For instance, the observables used in the b-jet tagging could be exploited in order to increase the discriminating power of the neural net. Several b-jet identification categories could also be studied. Some of them may also be used as an input to the network such as the b-jets containing a soft electron or muon [96]. Another multivariate approaches such as boosted decision trees could be tested in upcoming tagging studies.

In the run-I flavour tagging studies at the CMS experiment a lot of effort was invested in the building of the tagging framework. The re-weighting procedure of the simulated events as well as the reconstruction of $B^+ \to J/\psi K^+$ and $B_d^0 \to J/\psi K^{0*}$ events from the data and their use as a reference modes for the flavour tagging measurements were established. Therefore, even more emphasis could be put on the selection and optimisation of the tagging algorithms in the run-II tagging development as the complete tagging infrastructure now exists in the CMS experiment.

7. Weak phase and decay width difference measurement

In this chapter, we shortly summarise the results of the weak mixing phase analysis [12, 97] that utilises the MLP algorithm for flavour tagging. A complete description of the ϕ_s measurement can be found in Reference [75]. No details of the event reconstruction is given here, as the B_s^0 reconstruction and signal selection were already described in Section 6.3. The dataset used in the ϕ_s measurement is the same than the one used in the flavour tagging. The data were collected with a center-of-mass energy of 8 TeV, and correspond to an integrated luminosity of 19.7 fb^{-1}.

7.1 Maximum likelihood fit

A multidimensional unbinned extended maximum-likelihood fit to the data is performed to distributions of seven observables: B_s^0 invariant mass, three decay angles $\Theta = (\phi_T, \cos\theta_T, \cos\psi_T)$, flavour tag decision ξ, as well as ct and its uncertainty σ_{ct}. The likelihood function is

$$\begin{aligned} L &= N_{Sig} L_{Sig} + N_{Bkg} L_{Bkg} \\ L_{Bkg} &= P_{Bkg}(\cos\theta_T, \varphi_T) \cdot P_{Bkg}(\cos\psi_T) \cdot P_{Bkg}(ct) \cdot P_{Bkg}(m_{B_s}) \cdot P_{Bkg}(\sigma_{ct}) \cdot P_{Bkg}(\xi) \\ L_{Sig} &= \left(\tilde{f}(\Theta, \alpha, ct) \otimes G(ct, \sigma_{ct}) \cdot \epsilon(\Theta)\right) \cdot P_{Sig}(m_{B_s}) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi) \end{aligned} \tag{7.1}$$

where L_{Sig} is the likelihood model for the $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ signal component and L_{Bkg} is the likelihood function for the background. Parameters N_{Sig} and N_{Bkg} are the numbers of signal and background events, respectively. Function $\tilde{f}(\Theta, \alpha, ct)$ is the differential decay rate of Equation (2.111). The decay width difference $\Delta\Gamma_s$ is required to be positive in the fit following the measurement by the LHCb experiment [98]. Moreover, a Gaussian constraint is set to Δm_s such that the parameter is allowed to vary within one standard deviation of its world-average value $\Delta m_s = (17.69 \pm 0.08) \cdot 10^{-12}$ $\hbar$/s [99]. The magnitude of the CP violating parameter $|\lambda_{\mathrm{J}/\psi\phi}|$ is set to unity. The strong phase δ_0 is set to zero in the likelihood function. The difference in phases $\delta_S - \delta_\perp$ is fitted with a single parameter $\delta_{S\perp}$ in order to decrease the correlations between the various free parameters of the fit.

The $G(ct, \sigma_{ct})$ is a Gaussian resolution function whose width is the ct uncertainty σ_{ct} determined on an event-by-event basis. The $\epsilon(\Theta)$ is the angular efficiency function that corrects for the reconstruction and selection effects in the angular distributions. The signal mass model $P_{Sig}(m_{B_s})$ is described by a sum of three Gaussian functions with a common mean. The background mass distribution is modelled with an exponential function $P_{Bkg}(m_{B_s})$. The background component of the ct distribution is parametrised using a sum of two exponential functions and the pdf is denoted as $P_{Bkg}(ct)$. The angular background distributions of $\cos\theta_T$ and $\cos\psi_T$ are described with a series of Legendre polynomials $P_{Bkg}(\cos\psi_T)$ and $P_{Bkg}(\cos\theta_T, \varphi_T)$, respectively, while sinusoidal functions are used to model the distribution of φ_T.

The signal and background components for the ct uncertainty are modelled with Gamma functions, $P_{Sig}(\sigma_{ct})$ and $P_{Bkg}(\sigma_{ct})$. The signal component is described with a sum of two Gamma functions, while one Gamma function is used to model the background. The tag decision pdfs for signal and background components, $P_{Sig}(\xi)$ and $P_{Bkg}(\xi)$, are obtained directly from the data.

7.2 Results and systematic uncertainties

The fit results and their statistical uncertainties are reported in Table 7.1, and the fit projections onto various B_s^0 observables are shown in Figure 7.1. The likelihood profiles of $\delta_{S,\perp}$, $\delta_{\parallel}$ and $|A_S|^2$ are not parabolic and hence the statistical uncertainties for these parameters are taken from the increase of the negative log-likelihood by 0.5. The CMS measurements with the 68% confidence-level contours and the results from the ATLAS, CDF, D0 and LHCb experiments are shown in the $\Delta\Gamma_s$ - ϕ_s plane in Figure 7.2 [100]. The CMS results are in agreement with the standard model predictions, $\phi_s = -0.0370 \pm 0.0006$ rad [4] and $\Delta\Gamma_s = 0.088 \pm 0.020$ ps^{-1} [101], and with the measurements done by other experiments. The combined results, $\phi_s = -0.021 \pm 0.031$ rad and $\Delta\Gamma_s = 0.090 \pm 0.005$ ps^{-1} [100], are also consistent with the SM predictions.

Systematic uncertainties for $\Delta\Gamma_s$ and ϕ_s are evaluated by testing the assumptions made in the model pdf and those related to the fit procedure. The leading systematic uncertainty of $\Delta\Gamma_s$ result arises from the *ct* efficiency. The largest contributions of uncertainties for ϕ_s originate from modelling of the angular efficiencies, likelihood model, and setting $|\lambda_{J/\psi\phi}|$ to unity in the fit pdf. Additionally, the uncertainty associated with small differences between the simulated and observed kaon p_T spectra has a non-negligible contribution in the total systematic uncertainty of the weak phase.

Several other systematic uncertainties are also evaluated but their contribution to the total uncertainties of $\Delta\Gamma_s$ and ϕ_s is shown to be small. The effects of the uncertainties in the mistag fraction evaluation and modelling of the wrong tag fraction distributions for signal and background components are quantified. The effect of changing the model pdfs is also tested. Finally, an effect of tracker misalignment is quantified. All the systematic uncertainties are listed in Table 7.2. As the last cross-check, the constraint on Δm_s is released and the parameter is let vary in the

fit. The best fit value of Δm_s is found to be in agreement with the world-average value.

	Fit result
ϕ_s	-0.075 ± 0.097 rad
$\Delta\Gamma_\mathrm{s}$	0.095 ± 0.013 ps^{-1}
$\lvert A_0\rvert^2$	0.510 ± 0.005
$\lvert A_S\rvert^2$	$0.012\,^{+0.009}_{-0.007}$
$\lvert A_\perp\rvert^2$	0.243 ± 0.008
$\delta_\parallel$	$3.48\,^{+0.07}_{-0.09}$ rad
$\delta_{S\perp}$	$0.37\,^{+0.28}_{-0.12}$ rad
$\delta_\perp$	2.98 ± 0.36 rad
$c\tau$	447.2 ± 2.9 μm
N_{sig}	49183 ± 260
N_{bkg}	21276 ± 198

Table 7.1: Results of the multidimensional fit to the $\mathrm{B_s^0}$ observables. The uncertainties are statistical only.

7.3 Summary and outlook of the ϕ_s analysis

The weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ were measured using the data collected by the CMS detector at a center-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.7 fb^{-1}. In a total of 49 200 $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ signal candidates were extracted from the data. The analysis used opposite-side electron and muon tagging to identify the $\mathrm{B_s^0}$ flavour at its production time. The results for the weak mixing phase and the decay width difference are $\phi_\mathrm{s} = -0.075 \pm 0.097$ (stat) $\pm$ 0.031 (syst) rad and $\Delta\Gamma_\mathrm{s} = 0.095 \pm 0.013$ (stat) $\pm$ 0.007 (syst) ps^{-1}, respectively. The measured values are in agreement with the results from other experiments and

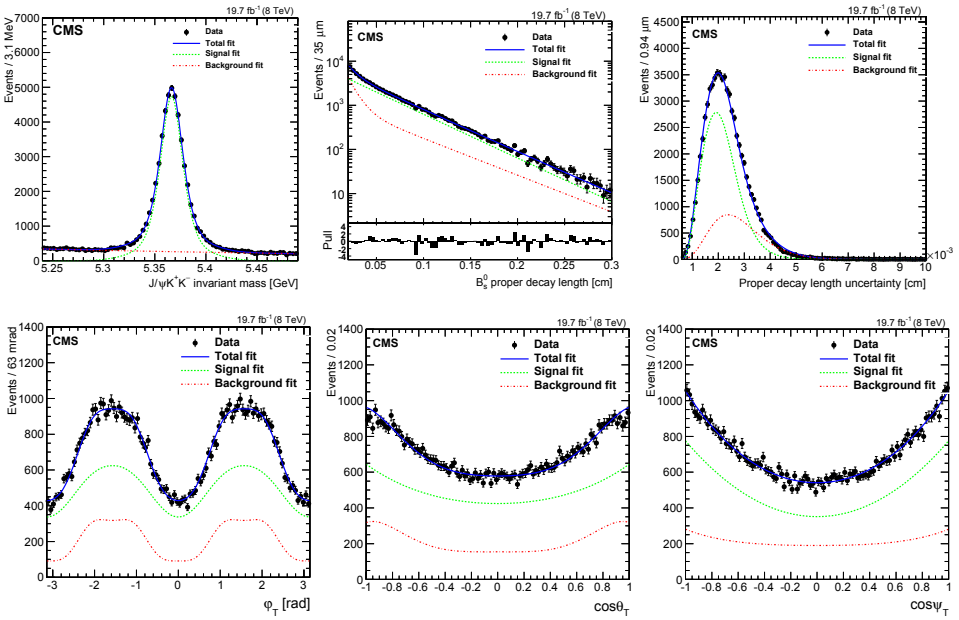


Figure 7.1: Fit projections onto the six observables of the B_s^0 system including mass, ct, and its uncertainty and the three angular distributions $\phi_T, \cos\theta_T$, and $\cos\psi_T$.

the SM predictions. The CMS measurements improve the combined results of these quantities and thus contribute to probe the standard model. The precision of our measurements is dominated by the statistical uncertainty. However, a sizeable reduction in statistical uncertainties is foreseen when performing the analysis using a new, larger dataset collected with a center-of-mass energy of 13 TeV.

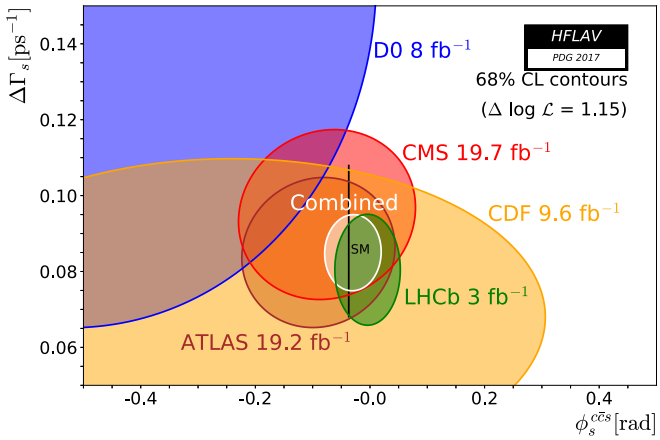


Figure 7.2: The individual 68% confidence-level contours of ATLAS, CMS, CDF, D0 and LHCb results in $\Delta\Gamma_s$ - ϕ_s plane. The combined contour is denoted as a solid white line and a shaded area. The standard model predictions for $\Delta\Gamma_s$ and ϕ_s are marked with a thin black rectangle. Figure is taken from [100].

Source of uncertainty	ϕ_s [rad]	$\Delta\Gamma_s$ [ps^{-1}]	$\lvert A_0\rvert^2$	$\lvert A_S\rvert^2$	$\lvert A_\perp\rvert^2$	$\delta_\parallel$ [rad]	$\delta_{S\perp}$ [rad]	$\delta_\perp$ [rad]	$c\tau$ [μm]
ct efficiency	0.002	0.0057	0.0015	-	0.0023	-	-	-	1.0
Angular efficiency	0.016	0.0021	0.0060	0.008	0.0104	0.674	0.14	0.66	0.8
Kaon p_T weighting	0.014	0.0015	0.0094	0.020	0.0041	0.085	0.11	0.02	1.1
ct resolution	0.006	0.0021	0.0009	-	0.0008	0.004	-	0.02	2.9
Mistag distribution modelling	0.004	0.0003	0.0006	-	-	0.008	0.01	-	0.1
Flavour tagging	0.003	0.0003	-	-	-	0.006	0.02	-	-
Model bias	0.015	0.0012	0.0008	-	-	0.025	0.03	-	0.4
Modelling of distributions	0.006	0.0021	0.0016	0.002	0.0021	0.010	0.03	0.04	0.2
$\lvert\lambda\rvert$ as a free parameter	0.015	0.0003	0.0001	0.005	0.0001	0.002	0.01	0.03	-
Tracker alignment	-	-	-	-	-	-	-	-	1.5
Total systematic uncertainty	0.031	0.0070	0.0114	0.022	0.0116	0.680	0.18	0.66	3.7
Statistical uncertainty	0.097	0.0134	0.0053	0.008	0.0075	0.081	0.17	0.36	2.9

Table 7.2: Systematic uncertainties of the various fit parameters. If the value of the uncertainty is not given, the uncertainty is negligible compared to the statistical and other systematic uncertainties. The total systematic uncertainty is the quadratic sum of the individual systematic uncertainties.

8. Measurement of the effective lifetime in the $\mathrm{B_S^0} \to \mathrm{J}/\psi\phi$ decay

This chapter reviews the effective lifetime measurement of the $\mathrm{B_s^0}$ meson decaying into a $\mathrm{J}/\psi\phi$ state. The effective lifetime analysis is done in the context of the CMS run-I paper on the b-hadron lifetimes. The results of the lifetime measurements can be found in Reference [102]. In the lifetime paper, the $\mathrm{B^+} \to \mathrm{J}/\psi\mathrm{K^+}$ decay is used as a reference mode for the various cross-checks and studies of the systematic uncertainties. Those cross-checks are not discussed here in detail, but they are mentioned in the section of systematic uncertainties. In addition, the systematic uncertainty of the tracker misalignment is evaluated using $\mathrm{B_d^0} \to \mathrm{J}/\psi\mathrm{K}^{0*}$ decay and generalised to other decays that have the same topology. The studies related to the $\mathrm{B^+}$ and $\mathrm{B_d^0}$ decays are done by Cinvestav group in Mexico and are briefly described here.

8.1 Feasibility studies

Since the decay width difference $\Delta\Gamma_s$ is large in the $\mathrm{B_s^0}$ system, it is not evident that the effective lifetime can be accurately measured by fitting a single exponential function to the *ct* distribution. Both upper and lower limits of the fit range may affect the effective lifetime, as the admixture of the heavy and light states in the *ct* distribution may change due to these cuts. In this analysis, the lower limit of the

ct fit range is set to 200 μm to remove resolution and efficiency effects present in the low-ct region. In order to ensure that the fit gives an expected result, a set of pseudo-experiments are done with and without cutting the ct distribution at 200 μm and testing the effect of the upper limit of the fit range.

8.1.1 Upper limit of the fit range

The effect of the upper limit ct cut on the effective lifetime is tested using about 4300 pseudo-experiments. The ct distributions are generated using the decay rate in Equation (2.90) and fitted with an exponential function in the ranges of [0,10000] μm (full range), [0,6000] μm, [0,4500] μm, and [0,3000] μm. Each ct distribution contains $20 \cdot 10^6$ events. The prediction for the effective lifetime ($c\tau_{gen}$) are obtained from Equations (2.92)- (2.94) and replacing the upper limit of the integrals by the above mentioned values. The theoretical lifetime in each fit range is within 441.73-441.74 μm. Based on the pseudo-experiments, pull distributions, $(c\tau_{fit} - c\tau_{gen})/\Delta c\tau_{fit}$, are formed. The fractions of events excluded from the fit ranges are also quantified. The values used for the lifetimes and amplitudes are shown in Table 8.1. The pull distributions are shown in Figure 8.1, and the key statistics of lifetime and pull distributions are reported in Table 8.2.

	world-average (2016) values
$c\tau_H$ [μm]	482.7
$c\tau_L$ [μm]	426.3
$\|A_\perp\|^2$	0.25
$\|A_\parallel\|^2 + \|A_0\|^2$	0.75

Table 8.1: The physics parameters used in the pseudo-experiments related to the upper limit of the fit range. The lifetimes and amplitudes are set to their world-average values [2].

Pseudo-experiments show a bias in the lifetime measurement in the fit ranges

	[0,0.3] cm	[0,0.45] cm	[0,0.6] cm	[0,1] cm
mean of the fitted lifetimes (μm)	441.535 ± 0.002	441.717 ± 0.002	441.741 ± 0.002	441.744 ± 0.002
sigma of the fitted lifetimes (μm)	0.101 ± 0.001	0.099 ± 0.001	0.099 ± 0.001	0.099 ± 0.001
mean of the pull distribution	-1.922 ± 0.015	-0.249 ± 0.015	-0.012 ± 0.015	0.012 ± 0.015
sigma of the pull distribution	1.000 ± 0.011	1.005 ± 0.011	1.005 ± 0.011	1.005 ± 0.011
mean fraction of excluded events	$1.2 \cdot 10^{-3}$	$4.4 \cdot 10^{-5}$	$1.7 \cdot 10^{-6}$	0

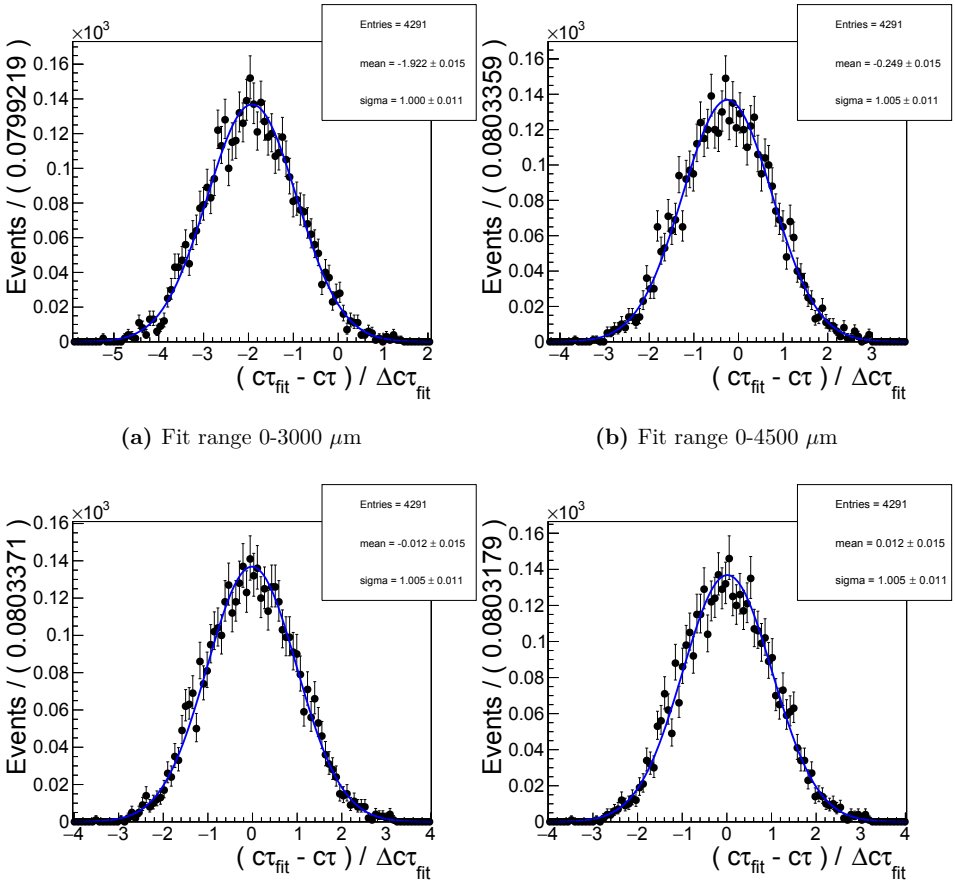
Table 8.2: Key statistics of the lifetime and pull distributions of the maximum likelihood fits.

of [0,3000] μm and [0,4500] μm. Theoretical predictions for effective lifetimes with varying upper limits of the fit range are, however, practically the same: 441.73 - 441.74 μm. Thus, the upper limit of the fit does not change the fractions of heavy and light states in the sample. The bias may originate from the fit artifact related to the fact that an exponential function is fitted to a double exponential decay rate distribution.

The bias increases when the fit range decreases since more events are excluded from the distribution. The biggest bias of -0.2 μm results from fitting the range 0-3000 μm. The fit in the range of 0-4500 μm shows a small negative bias of 0.025 μm. Since the fit interval 0-6000 μm excludes only a few tens of events in the large statistics pseudo-experiments, no bias is present in the fit results obtained from this range. The upper limit of 6000 μm is hence used in the effective lifetime measurement.

8.1.2 Lower limit of the fit range

The lower limit of the fit range is tested using about 5000 pseudo-experiments containing $2 \cdot 10^6$ events each. The generated ct distributions are fitted with an exponential function with and without cutting at $ct = 200$ μm. The upper limit of the fit range is set to 6000 μm, as discussed earlier. The lifetimes of the B_s^0 eigenstates are set to the values used in one of the simulated $B_s^0 \to J/\psi\phi$ samples. The world-



(a) Fit range 0-3000 μm

(b) Fit range 0-4500 μm

(c) Fit range 0-6000 μm

(d) Fit range 0-1 cm

Figure 8.1: Pull distributions of an exponential fit where the upper limit of the fit range is varied.

average values [103] are also used. The values of the physics parameters are listed in Table 8.3.

The predictions for the effective lifetimes ($c\tau_{gen}$) are obtained using Equations (2.94) and (2.97) and reported in Table 8.4. The pull distributions are shown in Figure 8.2. The key statistics of the pull distributions is listed in Tables 8.5 - 8.6. No bias is observed in the mean of the pulls. The uncertainties of the fits are also estimated correctly as the widths of the pull distributions are consistent with unity.

	MC parameters	world-average (2014) parameters
$c\tau_H$ [μm]	474.6	498.0
$c\tau_L$ [μm]	409.8	421.2
$\|A_\perp\|^2$	0.246	0.246
$\|A_\parallel\|^2 + \|A_0\|^2$	0.754	0.754

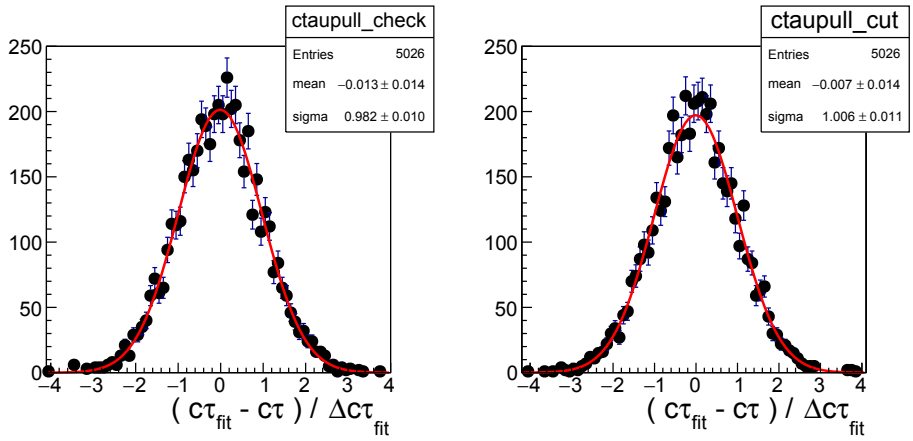
Table 8.3: The physics parameters used in the pseudo-experiments. The lifetimes are taken from one of the simulated B_s^0 samples and set to their world-average values [103]. The amplitudes are set to the values used the simulated $B_s^0 \to J/\psi\phi$ sample.

8.2 Data and simulated samples

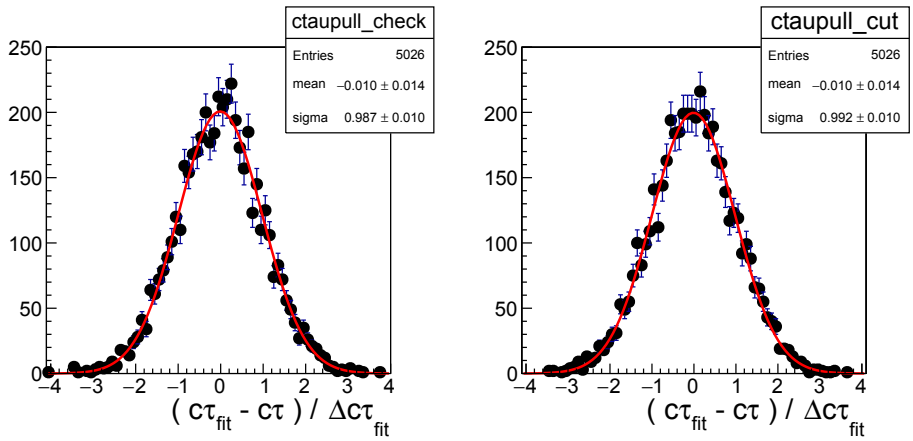
The data used in this analysis were collected with a center-of-mass energy of 8 TeV during run I of the LHC, and correspond to an integrated luminosity of 19.7 fb^{-1}. The analysis exploits a large number of simulated samples, including B_s^0 and other

	$ct > 0$ μm	$ct > 200$ μm
MC $c\tau_{eff}$ [μm]	427.55	428.42
world-average $c\tau_{eff}$ [μm]	442.57	443.71

Table 8.4: Predictions for effective lifetimes in different values for lifetimes and different ct cut values.



(a) Pulls from the generation of events using the MC lifetimes: Fits are performed in the full *ct* range 0-6000 μm (left) and in the nominal range 200-6000 μm (right).



(b) Pulls from the generation of events using the world-average lifetimes: Fits are performed in the full *ct* range 0-6000 μm (left) and in the nominal range 200-6000 μm (right).

Figure 8.2: Pull distributions used to study feasibility of the effective lifetime measurement.

	[0,6000] μm	[200,6000] μm
mean of the pull distribution	-0.014 ± 0.014	-0.007 ± 0.014
sigma of the pull distribution	0.982 ± 0.010	1.006 ± 0.011

Table 8.5: Key statistics of the pull distributions. The MC lifetimes are used in the generation of the pseudo-experiments.

	[0,6000] μm	[200,6000] μm
mean of the pull distribution	-0.010 ± 0.014	-0.010 ± 0.014
sigma of the pull distribution	0.987 ± 0.010	0.992 ± 0.010

Table 8.6: Key statistics of the pull distributions. The world-average values of the lifetimes are used in the generation of the pseudo-experiments.

b-hadron samples. The simulated samples used in the effective lifetime measurement are listed below, and a short description of their application is given.

- Simulated exclusive $B_s^0 \to J/\psi\phi$ sample with decay width difference $\Delta\Gamma_s = 0$ ps^{-1}
  - The sample is used to evaluate the effect of reconstruction and selection efficiencies affecting the reconstructed B_s^0 *ct* distribution.
- Simulated exclusive $B_s^0 \to J/\psi\phi$ sample with decay width difference $\Delta\Gamma_s = 0.1$ ps^{-1}
  - The sample is used to evaluate the total efficiency affecting to the reconstructed B_s^0 *ct* distribution. The sample is also utilised in the validation of the likelihood function and the fit procedure.
- Simulated inclusive $B_s^0 \to J/\psi(\mu\mu)X$ sample with decay width difference $\Delta\Gamma_s = 0.068$ ps^{-1}

  - The sample is used to evaluate the fraction of the S-wave component in the reconstructed $B_s^0 \to J/\psi\phi$ events. Furthermore, this sample is used in the background composition analysis of the B_s^0 mass distribution.

- Other simulated inclusive b hadron samples $B_d^0 \to J/\psi(\mu\mu)X$, $B^+ \to J/\psi(\mu\mu)X$ and $\Lambda_b \to J/\psi(\mu\mu)X$

  - The samples are used in the background composition analysis of the B_s^0 mass distributions.

The simulated samples have been generated over the years 2012-2014 and contain different values for the physics parameters Γ_s,$\Delta\Gamma_s$, $|A_i|^2$ and δ_i. The parameter values used in the B_s^0 samples, together with their world-averages, are listed in Table 8.7. Data-simulation comparisons of the key observables constructing the B_s^0 ct, and ct uncertainty distributions are shown in Appendix A.

	1^{st} excl. $B_s^0 \to J/\psi\phi$ MC	1^{st} incl. B_s^0 MC	2^{nd} excl. $B_s^0 \to J/\psi\phi$ MC	world-average (2016)
$c\tau$	439.8 μm	438.0 μm	438.0 μm	452.7 ± 1.5 μm
$\Delta\Gamma_s$	0.1 ps^{-1}	0.068 ps^{-1}	0	0.082 ± 0.007 ps^{-1}
$\frac{dG}{\Gamma_s}$	14.7%	10%	0 %	12.4 ± 1.1%
β_s $(-\frac{1}{2}\phi_s)$	0.02 rad	0.02 rad	0.02 rad	0.006 ± 0.019 rad
Δm_s	17.69 ps^{-1}	20.0 ps^{-1}	20.0 ps^{-1}	17.757 ± 0.021 ps^{-1}
$\lvert\lambda\rvert$	1	1	1	1.02 ± 0.07
δ_0	0 rad	0 rad	0 rad	0 rad
$\delta_\perp$	3.14 rad	-0.17 rad	-0.17 rad	3.16 ± 0.24 rad
$\delta_\parallel$	3.14 rad	2.50 rad	2.50 rad	$3.23^{+0.10}_{-0.14}$ rad
$\lvert A_\perp\rvert^2$	0.2460	0.16	0.16	0.250 ± 0.006
$\lvert A_\parallel\rvert^2$	0.2313	0.24	0.24	0.224 ± 0.010
$\lvert A_0\rvert^2$	0.5227	0.60	0.60	0.528 ± 0.006

Table 8.7: Physics parameters used in the three simulated B_s^0 samples and their corresponding world-average values.

8.3 Event reconstruction and selection

The event reconstruction and signal selection are based on the same procedures applied in the flavour tagging and weak mixing phase studies [12, 82, 83]. However, the trigger is switched from a displaced J/ψ trigger to a simplified dimuon trigger without a displacement requirement. The new trigger has less stringent selection imposed on the muons and J/ψ meson compared to the displaced J/ψ trigger. The trigger is changed in order to obtain larger yield of B_s^0 signal events compared to the weak phase analysis. Absence of the displacement requirement at the trigger level also reduces the possible biases affecting the lifetime measurement. In addition, the χ^2 vertex fit probability requirement for the B_s^0 mesons is loosened from the previous selection to increase the number of signal events.

8.3.1 Trigger selection

The data are collected with a high-level trigger optimised for detecting a J/ψ meson. The trigger reconstructs a J/ψ meson from an oppositely-charged muon pair using a set of loose muon selection criteria. The muons with $|\eta| < 2.2$ are selected and the muon tracks are required to have the distance of closest approach less than 0.5 cm. The muon tracks are fitted to a common vertex. The χ^2 vertex fit probability is required to be at least 0.5%. The invariant mass of J/ψ candidate must be within 2.9-3.3 GeV, and the p_T of the J/ψ candidate has to be greater than 7.9 GeV.

8.3.2 Muons and J/ψ mesons

In the offline analysis the muons have to be classified as loose muons and matched to the muons detected by the dimuon trigger. The muons are required to originate from a common vertex whose position and validity is determined using the Kalman vertex fit. The other offline cuts for the muon properties such as p_T, $|\eta|$ and DCA are set as

tight as in the trigger selection. The J/ψ mesons are reconstructed from oppositely charged muon pairs. The mass of the J/ψ candidate is required to be within 150 MeV of the world-average mass [84]. The offline cuts for the J/ψ properties such as p_T and vertex probability are also required to be as tight as in the trigger selection.

8.3.3 $\phi(1020)$ mesons

The $\phi(1020) \rightarrow KK$ candidates are reconstructed from oppositely-charged tracks that have $p_T > 0.7$ GeV. Each track has to comprise a minimum of five tracker hits. Each track is assumed to be a kaon. Muon candidate tracks are removed from the collection in which the kaon tracks are selected. The invariant mass of the track pair is required to lie within 10 MeV of the world-average $\phi(1020)$ mass [84].

8.3.4 B_s^0 mesons

The B_s^0 mesons are reconstructed by performing a kinematic fit to the muon and kaon tracks that originate from the J/ψ and $\phi(1020)$ candidates. The dimuon mass is constrained to the J/ψ mass [84] in the kinematic fit. A B_s^0 candidate is retained if its invariant mass lies within the window of 5.24-5.5 GeV and its χ^2 vertex fit probability is greater than 0.1%. If the event has multiple B_s^0 candidates, the one having the highest vertex probability is selected. The ct of the B_s^0 candidate is also required to be greater than 200 μm.

8.4 Fit observables

The main observables used in the effective lifetime measurement are the mass m, proper decay time t times the speed of light, ct, and ct uncertainty σ_{ct}. The mass distribution of the B_s^0 candidate is used to extract the signal and background components of the ct and ct uncertainty distributions. The decay time t measured in

the transverse plane is defined as

$$t = m\frac{L_{xy}}{p_{\mathrm{T}}}, \tag{8.1}$$

where m is the mass of the $\mathrm{B_s^0}$ candidate set to the world-average value, L_{xy} is the transverse flight distance, and p_{T} is the transverse momentum. The transverse flight distance is defined as

$$L_{xy} = \frac{\boldsymbol{v}_{\mathrm{T}} \cdot \boldsymbol{p}_{\mathrm{T}}}{p_{\mathrm{T}}} = v_{\mathrm{T}} \cos(\boldsymbol{v}_{\mathrm{T}}, \boldsymbol{p}_{\mathrm{T}})$$
$$\boldsymbol{v}_{\mathrm{T}} = \begin{pmatrix} x_{SV} - x_{PV} \\ y_{SV} - y_{PV} \end{pmatrix},$$

where x_{PV}, x_{SV}, y_{PV} and y_{SV} are the x- and y-coordinates of primary (PV) and secondary (SV) vertices in the laboratory frame. The ct uncertainty is obtained using the error propagation but omitting correlations

$$\sigma_{ct} = \sqrt{\left(\frac{m}{p_{\mathrm{T}}}\sigma_{L_{xy}}\right)^2 + \left(-\frac{mL_{xy}}{p_{\mathrm{T}}^2}\sigma_{p_{\mathrm{T}}}\right)^2}, \tag{8.2}$$

where $\sigma_{p_{\mathrm{T}}}$ and $\sigma_{L_{xy}}$ are uncertainties of p_{T} and L_{xy}. The uncertainty of L_{xy} is defined as

$$\sigma_{L_{xy}} = \sigma_{v_{\mathrm{T}}} \cdot |\cos(\boldsymbol{v}_{\mathrm{T}}, \boldsymbol{p}_{\mathrm{T}})|$$
$$\sigma_{v_{\mathrm{T}}} = \sqrt{\frac{\boldsymbol{v}_{\mathrm{T}}^T E_{xy} \boldsymbol{v}_{\mathrm{T}}}{v_{\mathrm{T}}^2}}$$
$$E_{xy} = E_{xy}^{SV} + E_{xy}^{PV},$$

where E_{PV} and E_{SV} are the covariance matrices of primary and secondary vertices. The uncertainty associated with the $\mathrm{B_s^0}$ transverse momentum is assumed to be proportional to the covariance matrix of secondary vertex in such a way that

$$(L_{xy}\sigma_{p_T})^2 = \boldsymbol{v}_{\mathrm{T}}^T E_{xy}^{SV} \boldsymbol{v}_{\mathrm{T}} \left(\cos(\boldsymbol{v}_{\mathrm{T}}, \boldsymbol{p}_{\mathrm{T}})\right)^2.$$

8.5 Background studies

A set of simulated b-hadron samples is studied to make sure there are no peaking backgrounds in the B_s^0 signal region to bias the effective lifetime. The decays of $B_s^0 \to J/\psi(\mu\mu)X$, $B_d^0 \to J/\psi(\mu\mu)X$, $B^+ \to J/\psi(\mu\mu)X$, and $\Lambda_b \to J/\psi(\mu\mu)X$ are analysed using the B_s^0 reconstruction and selection algorithms. The B_s^0 mass distribution is reconstructed from the b-hadron events that pass the signal selection. The events passing the selection are scaled to correspond to the total number of signal and background events in the data. No peaking background is observed in the B_s^0 mass region as shown in Figure 8.3.

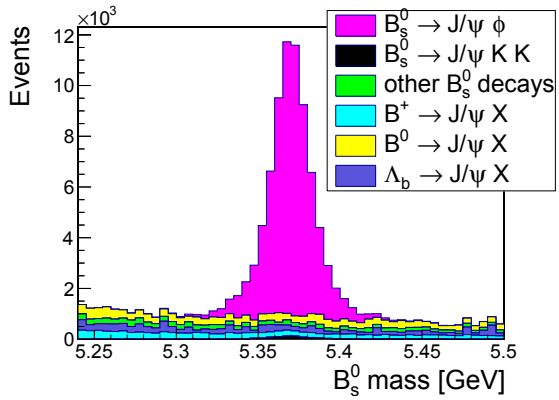


Figure 8.3: B_s^0 mass distribution constructed from inclusive simulated B_s^0, B^+, B_d^0 and Λ_b samples.

8.6 Vertex reconstruction and selection

In the 8 TeV energy, about 22 primary interactions occur during a LHC bunch crossing in the CMS detector. One of the interaction points is selected as the production vertex of the B_s^0 meson. Three methods to determine the primary vertex (PV) candidate are studied, and the position resolutions of the PV candidates are

quantified. The primary vertex candidates are selected based on

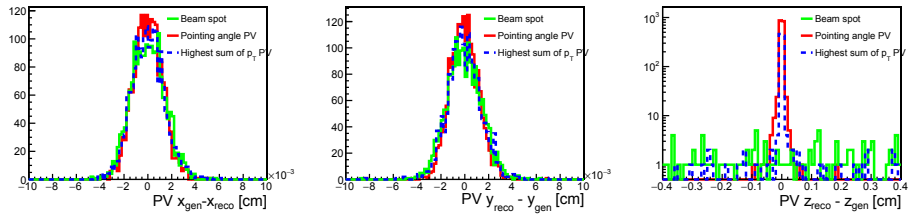
1. beam spot
2. the smallest pointing angle between the B_s^0 flight direction and the momentum
3. the highest sum of p_T of the tracks originating from a PV

The primary vertices fitted from the reconstructed tracks using the beam spot position as a constraint. The muon and kaon tracks arising from the decay of the B_s^0 candidate are removed from the PV fit.

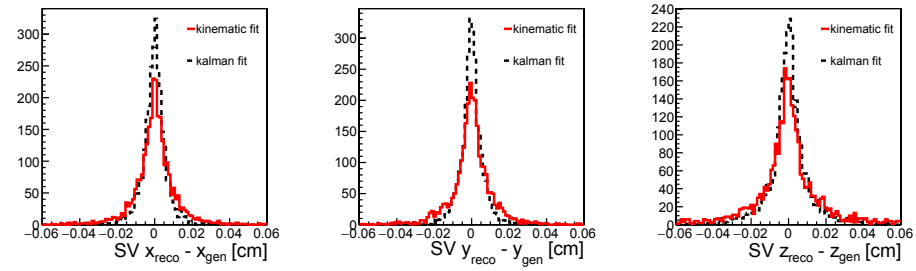
Two reconstruction techniques for the secondary vertices (SV) are also studied. The SVs are formed with a Kalman filter technique and using a kinematic fit.

The residual differences of the reconstructed primary and secondary vertex coordinates with respect to the generated ones are determined from the simulated exclusive B_s^0 sample with $\Delta\Gamma_s = 0.1$ ps^{-1}. The widths of these residual distributions are considered as resolutions for the PV and SV coordinates. The residuals are shown in Figure 8.4.

The residual distributions of the x- and y-coordinates have a similar width, while the width of the z-coordinate is about two orders of magnitude bigger. The secondary vertex resolution from the kinematic fit is worse compared to the Kalman filter approach, as can be seen in Figure 8.4. Thus, the Kalman filter technique is used to obtain the secondary vertex position. It is also apparent that the residual in the z-coordinate has a wider spread compared to the ones in the x- and y-coordinates. This means that a better resolution for the ct observable can be obtained in transverse plane compared to the full 3D reconstruction. Hence, the 2D ct distribution is used in the effective lifetime measurement.



(a)



(b)

Figure 8.4: Distributions of residual difference for a) primary vertex and b) secondary vertex coordinates.

8.7 Vertex and ct resolutions

The vertex position and ct resolutions are determined from the residual distributions of vertex coordinates using the simulated exclusive $B_s^0 \to J/\psi\phi$ sample. The resolutions are obtained by fitting either a Gaussian or a double Gaussian function to each of the distributions. For a single Gaussian fit, the resolution is obtained from the width of the Gaussian. For the double Gaussian fit functions, resolution is defined as

$$\sigma_{res} = \sqrt{f_1\sigma_1^2 + f_2\sigma_2^2}, \tag{8.3}$$

where the parameters f_i are the relative strengths between the Gaussians and σ_i are the standard deviations of the Gaussians pdfs.

The vertex position resolution for the primary vertices measured with different

selection is about 12-14 μm for the x- and y-coordinates. The resolution for the x- and y-coordinates of the secondary vertex is about 60 μm when the Kalman filter approach is used and 96 μm when the vertex position is determined with a kinematic fit. The position resolutions for the primary and secondary vertices are summarised in Table 8.8. Although the resolutions are similar for all the techniques to reconstruct a PV, the pointing angle method is used in the effective lifetime measurement. It was also used in the weak mixing phase analysis [12]. The ct resolution is 25 μm. The resolutions obtained using different PV reconstruction methods are reported in Table 8.9. The primary and secondary vertex resolutions found in this analysis are in agreement with the results reported in Chapter 5, Table 5.1.

	σ_x [μm]	σ_y [μm]
PV, highest sum of track p_{T}	13.4	13.5
PV, pointing angle	12.2	11.8
Beam spot	13.9	14.3
SV (kalman)	60.0	58.1
SV (kinematic fit)	96.0	95.8

Table 8.8: The PV and SV position resolutions for the x- and y-coordinates.

	σ_{ct} [μm]
PV, highest sum of track p_{T}	24.8
PV, pointing angle	24.6
Beam spot	24.6

Table 8.9: The ct resolution determined for the primary vertices chosen with different selection criteria.

8.8 ct efficiency

The reconstructed ct distribution is affected by the detector acceptance and inefficiencies in the reconstruction and selection of the B_s^0 events. These effects have to be corrected for to obtain an unbiased effective lifetime. The ct efficiency consists of two independent efficiencies, reconstruction efficiency $\epsilon_{reco}(ct)$ and apparent efficiency $\epsilon_{app}(ct)$.

8.8.1 Apparent efficiency

The B_s^0 meson system can be described by the following three features:

1. B_s^0 meson has two mass eigenstates that have differing masses and lifetimes
2. The mass eigenstates coincide with the CP eigenstates, since the CP violation in mixing is observed to be small
3. CP eigenstates are characterised by differing angular distributions

Based on the points 2) and 3) it is evident that the decays of the mass eigenstates are characterised by different angular distributions. The mass eigenstates are further described by the different lifetimes. The theoretical projections of the angular and ct distributions of the heavy and light mass eigenstates are shown in Figure 8.5.

As a consequence of different lifetimes and angular distributions of the mass eigenstates, the shapes of the angular distributions evolve as a function of the B_s^0 proper time. Moreover, the shapes of the angular distributions reflect the changes in the fraction of heavy and light states in the B_s^0 sample. A hypothetical example of the time-evolution of B_s^0 eigenstates is shown in Figure 8.6. The time-evolution of the fraction f_H of the heavy mass state is illustrated in Figure 8.6 and can be seen by an altering shape of the $\cos\theta$ distribution.

Angular distributions are distorted by a non-flat three-dimensional angular efficiency. Due to the time evolution of the B_s^0 system, the efficiency affects the

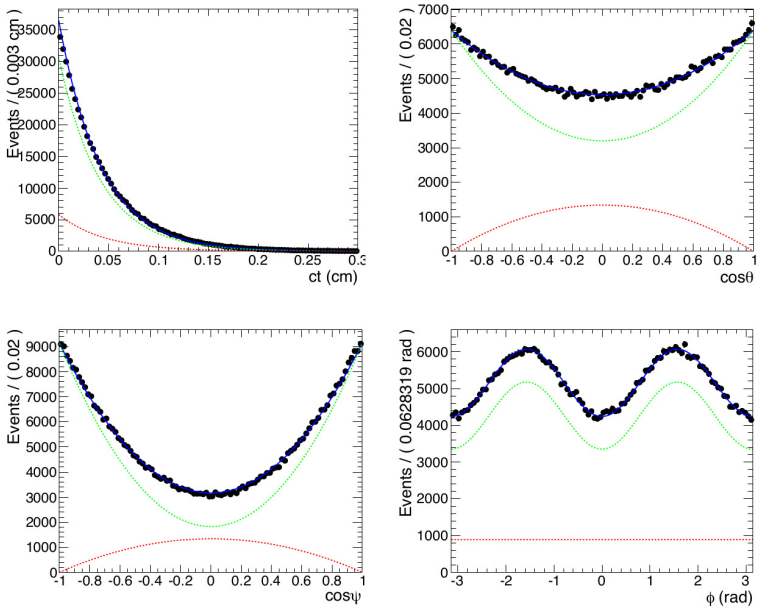
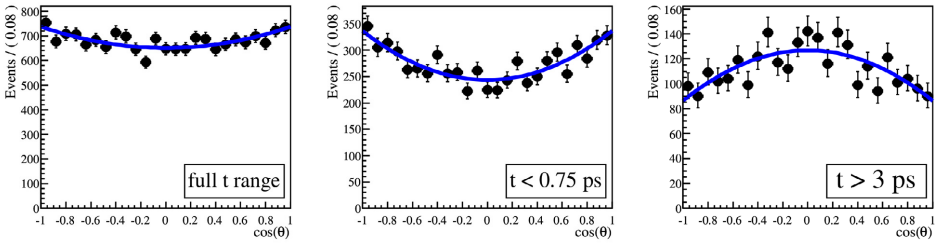


Figure 8.5: The theoretical projections of the angular and ct distributions of the heavy and light B_s^0 mass eigenstates shown in dashed red and green lines, respectively. Total distributions are shown in solid blue line. Figure is taken from [82].



(a)

Figure 8.6: Time evolution of $\cos\theta$ distribution. At short decay times light state dominates $\cos\theta$ distribution, as seen by the shape of the distribution. Going towards larger decay times heavy state starts to dominate the $\cos\theta$ distribution. Figure is taken from [104].

angular distributions differently as a function of the B_s^0 proper time. This means that the angular efficiency excludes different numbers of heavy and light state events as a function of the decay time. The phenomenon changes the fractions of heavy and light states in the B_s^0 sample and is called an apparent efficiency [104]. Apparent efficiency can be thought of as the *efficiencies of the amplitudes* $\epsilon_i = \left\{\epsilon_\perp, \epsilon_\parallel, \epsilon_0, \epsilon_{\perp,0}..\right\}$ which are defined as

$$\epsilon_i = \int \epsilon(\Theta) g_i(\Theta) d\Theta, \tag{8.4}$$

where $\epsilon(\Theta)$ is the angular efficiency function and $g_i(\Theta)$ is the i-th component of the angular distributions shown in Table 2.2. Considering the six terms in the untagged decay rate function in Equation (2.89), the apparent efficiency is given by

$$\epsilon_{app}(ct) = \frac{\sum_i \epsilon_i O_{i,tot}(\alpha, ct)}{\sum_i \int O_{i,tot}(\alpha, ct) g_i(\Theta) d\Theta}. \tag{8.5}$$

This can be explicitly written as

$$\begin{aligned}\epsilon_{app}(ct) = &\frac{\epsilon_0|A_0|^2 + \epsilon_\parallel|A_\parallel|^2 + \epsilon_\perp|A_\perp|^2 e^{\Delta\Gamma_s ct} + \epsilon_{0,\parallel}\cos(\delta_\parallel - \delta_0)|A_0||A_\parallel|}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 e^{\Delta\Gamma_s ct}} - \\ &\frac{\frac{1}{2}\sin\phi_s|A_\perp|(e^{\Delta\Gamma_s ct} - 1)\left\{\epsilon_{\parallel,\perp}|A_\parallel|\cos(\delta_\perp - \delta_\parallel) + \epsilon_{\perp,0}|A_0|\cos(\delta_\perp - \delta_0)\right\}}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 e^{\Delta\Gamma_s ct}},\end{aligned} \tag{8.6}$$

by assuming that $\cos\phi_s \approx 1$ and $|\lambda_{J/\psi\phi}| = 1$. The apparent efficiency depends on the decay width difference $\Delta\Gamma_s$, the magnitudes of the amplitudes $|A_i|^2$ and the shape of the angular efficiency that determines the parameters ϵ_i. The apparent efficiency is determined from the decay rate models designed for the weak mixing phase measurement [75]. Using these models, the apparent efficiency becomes

$$\epsilon_{app}(ct) = \frac{\int \frac{d^4\Gamma_{gen}(ct)}{d\Theta dct}\epsilon(\Theta)d\Theta}{\int \frac{d^4\Gamma_{gen}(ct)}{d\Theta dct}d\Theta}, \tag{8.7}$$

where the numerator is the generator-level B_s^0 decay rate model multiplied with the angular efficiencies, and the denominator contains the pure generator-level model. The projections of the three-dimensional angular efficiency function $\epsilon(\Theta)$ onto three angular observables $\cos(\Psi)$, $\cos(\theta)$ and ϕ are shown in Figure 8.7.

Apparent efficiencies evaluated using different values of $\Delta\Gamma_s$ are shown in Figure 8.8. The values for the amplitudes, $\Delta\Gamma_s$ and the strong phases are set to the ones used in the exclusive simulated B_s^0 sample having $\Delta\Gamma_s = 0.1$ ps^{-1}. When $\Delta\Gamma_s$ is set to zero, the efficiency is flat.

Since the ϕ_s analysis is designed for a different trigger, signal selection, and dataset than those used in the effective lifetime measurement, apparent efficiency derived from the ϕ_s decay rate templates is not the same as the efficiency affecting this analysis. Even if the shape of the angular efficiency can be slightly different in the two analyses, one can still estimate its effect on the lifetime. Additionally, it is possible to quantify the difference in apparent efficiencies in data and simulations due to the difference of physics parameter values ($|A_i|^2$, Γ_s, $\Delta\Gamma_s$ and δ_i) used in the simulation and present in the data. The apparent efficiency present in the data is evaluated using the world-average values [2] for the physics parameters. The efficiency curves are shown in Figure 8.8.

Effect of apparent efficiency

The effect of apparent efficiency can be evaluated using a set of pseudo-experiments. In total, 4900 B_s^0 ct distributions having $2 \cdot 10^6$ events are generated from the theoretical B_s^0 decay rate model. The ct distributions are fitted using i) an exponential function and ii) an exponential function multiplied separately with two different apparent efficiency functions determined using world-average and MC values of the physics parameters. These efficiencies are shown in Figure 8.8 b). The mean values of the fitted lifetimes are listed in Table 8.10 and the distributions are shown in Figure 8.9.

The contribution of the apparent efficiency is defined as a residual difference of the results fitted with and without the efficiency functions. The apparent efficiency determined using the world-average values for physics parameters affect the

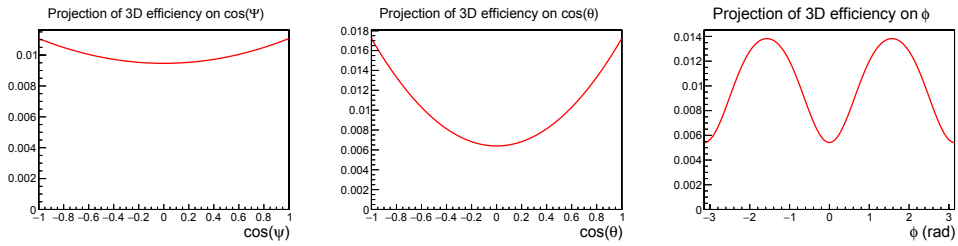
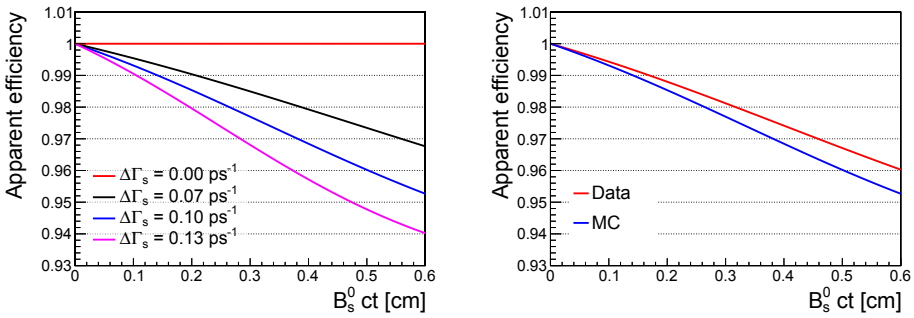


Figure 8.7: Fit projections of the 3D angular efficiency function used in the reconstruction level ϕ_s fit model.



(a) **(b)**

Figure 8.8: a) Apparent efficiencies evaluated for the simulated B_s^0 sample using different values of $\Delta\Gamma_s$. b) Apparent efficiencies evaluated for the simulated B_s^0 sample (blue) and for the data (red). Efficiencies are normalized to unity at $ct = 0.0$ cm.

	Mean of the distribution [μm]
fit without $\epsilon_{app}(ct)$	441.736 ± 0.005
fit using $\epsilon_{app}(ct)$ from data conf.	442.902 ± 0.005
fit using $\epsilon_{app}(ct)$ from MC conf.	443.153 ± 0.005

Table 8.10: The average values of the lifetime distributions obtained using an exponential fit pdf and multiplying it with an efficiency function.

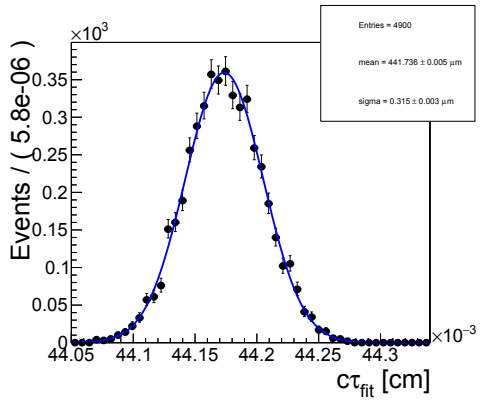
measured effective lifetime about 1.2 μm. The contribution of apparent efficiency obtained from the simulated sample is 1.4 μm. Thus, apparent efficiency affects the effective lifetime an order of a μm. The difference in physics parameters in the data and simulations does not have a big effect in the lifetime results. The difference in the lifetimes obtained using the two efficiencies is 0.25 μm.

8.8.2 Reconstruction efficiency

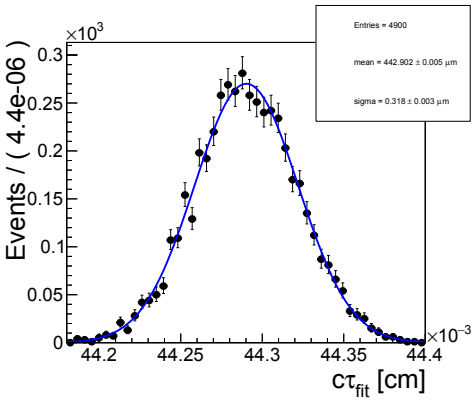
Reconstruction efficiency is the efficiency related to performance of the high-level trigger and the reconstruction of the muon and kaon tracks in the $B_s^0 \to J/\psi\phi$ decays. The reconstruction efficiency is obtained using a simulated generator-level $B_s^0 \to J/\psi\phi$ sample where $\Delta\Gamma_s$ is set to zero as the apparent efficiency is absent in this sample. The efficiency histogram is defined as

$$\epsilon(ct_i) = \frac{GEN_{after\,trigger+cuts}(ct_i)}{GEN(ct_i)}, \tag{8.8}$$

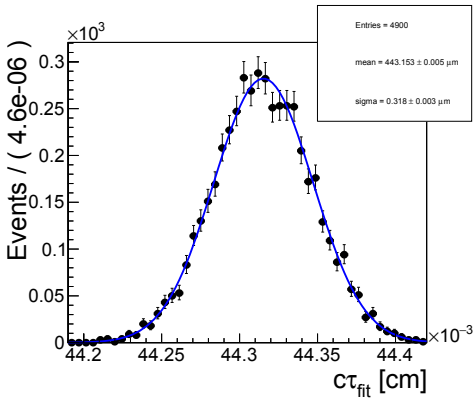
where the numerator is the generator-level ct histogram after reconstruction level trigger and selection cuts. The denominator $GEN(ct_i)$ is the pure generator-level ct histogram that is formed independently of the B_s^0 reconstruction. The parametrisation of the reconstruction and selection efficiency is presented in Equation (8.9), and the fitted efficiency function is shown in Figure 8.10. The fit results are reported in Table 8.11.



(a) Fit without $\epsilon_{app}(ct)$



(b) Fit using $\epsilon_{app}(ct)$ from data



(c) Fit using $\epsilon_{app}(ct)$ from MC

Figure 8.9: Pseudo-experiments performed by fitting a ct distribution with and without an apparent efficiency function.

$$\epsilon(ct) = k_0 + ct^{-k_1} + k_2/(ct^2 + k_3) \tag{8.9}$$

	Fit result
k_0	-0.777 ± 0.006
k_1	0.019 ± 0.002
k_2	$(-1.1 \pm 0.3) \cdot 10^{-5}$
k_3	0
χ^2/dof	26.4 / 31

Table 8.11: Fit parameters for the reconstruction efficiency from $\Delta\Gamma_s = 0$ sample, as well as χ^2 value and the number of degrees of freedom (dof) of the fit.

Effect of reconstruction efficiency

In order to quantify the effect of reconstruction efficiency in the lifetime result, the simulated sample having $\Delta\Gamma_s = 0$ ps^{-1} is fitted with and without the efficiency function. The results are compared to the generator-level value for effective lifetime. The fit results and the generator-level lifetime value are shown in Table 8.12. The fit results show that the contribution of the reconstruction efficiency is 10.4 μm. Moreover, the ct efficiency in the likelihood model corrects for the reconstruction and selection effects in the lifetime. The contribution of the reconstruction efficiency in the lifetime is an order of magnitude larger ($\sim$10 μm) compared to the apparent efficiency contribution ($\sim$1 μm).

8.8.3 Total ct efficiency

Total efficiency including the effects of apparent and reconstruction efficiencies is determined from the simulated exclusive $B_s^0 \to J/\psi\phi$ sample where $\Delta\Gamma_s = 0.1$ ps^{-1}.

	$c\tau_{eff}$ [μm]
generator-level value	438.0
fit without using $\epsilon_{reco}(ct)$	427.2 ± 1.1
fit using $\epsilon_{reco}(ct)$	437.6 ± 1.1

Table 8.12: Fit results with and without including the reconstruction efficiency in the model likelihood function.

The efficiency histogram is defined as

$$\epsilon(ct_i) = \frac{GEN_{after\,trigger+cuts}(ct_i)}{GEN_{\phi_s}(ct_i)}, \tag{8.10}$$

where the numerator is the generator-level ct histogram after reconstruction level trigger and selection cuts. The denominator $GEN_{\phi_s}(ct)$ is a ct distribution that is generated by the theoretical untagged decay rate model. The physics parameters of the theoretical model are set to those used in the simulated sample. The generator-level ct distribution is not used in the construction of the total efficiency histogram, since the loose selection on the track η and p_{T} is applied in the generator-level of the $\mathrm{B_s^0}$ sample, and this already induces the apparent efficiency in the generator-level ct distribution. The efficiency histogram is parametrised with the function shown in Equation (8.9) and fitted in the range 200-6000 μm to avoid the steep decrease of the efficiency due resolution and reconstruction effects present in the $ct < 200$ μm region. The total ct efficiency is shown in Figure 8.10, and the values of the fit parameters are reported in Table 8.13.

Effect of total efficiency

The effect of the total efficiency on the lifetime is estimated in a similar manner as the reconstruction efficiency. The likelihood function is fitted to the simulated $\mathrm{B_s^0}$ sample having $\Delta\Gamma_s = 0.1$ ps^{-1} with and without the ct efficiency function. The effective lifetimes obtained from the fits are presented in Table 8.14. The total

	Fit result
k_0	-0.891 ± 0.002
k_1	0.0039 ± 0.0006
k_2	$(9 \pm 2) \cdot 10^{-4}$
k_3	0.051 ± 0.007
χ^2/dof	28.2 / 31

Table 8.13: Fitted parameters of the total ct efficiency function, as well as χ^2 value and the number of degrees of freedom (dof) of the fit.

contribution of the efficiency to the lifetime is 12.1 μm, which is in agreement with the contributions of the reconstruction ($\sim$10 μm) and apparent ($\sim$1 μm) efficiencies discussed above.

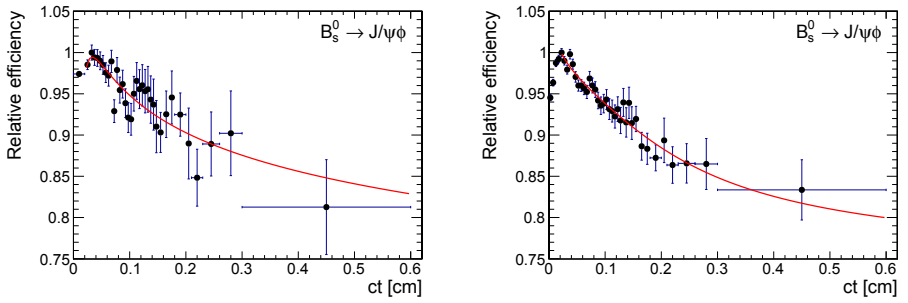
	$c\tau_{eff}$ [μm]
Predicted $c\tau_{eff}$	428.4
fit without using $\epsilon_{tot}(ct)$	416.5 ± 0.7
fit using $\epsilon_{tot}(ct)$	428.6 ± 0.7

Table 8.14: Fit results with and without including the total efficiency in the model likelihood function.

8.9 Likelihood function

The effective lifetime is obtained by performing a three-dimensional unbinned extended maximum-likelihood fit to distributions of reconstructed B_s^0 mass, ct, and ct uncertainty. The total likelihood function is

$$L_{tot} = N_{sig} L_{signal}(m, ct, \sigma_{ct}) + N_{bkg} L_{background}(m, ct, \sigma_{ct}), \tag{8.11}$$



(a) **(b)**

Figure 8.10: a) Reconstruction ct efficiency determined from the simulated $B_s^0 \to J/\psi\phi$ sample in which $\Delta\Gamma_s = 0$ ps^{-1}. b) Total ct efficiency that includes the reconstruction and apparent efficiencies. Total efficiency is obtained from the exclusice $B_s^0 \to J/\psi\phi$ sample in which $\Delta\Gamma_s = 0.1$ ps^{-1}. The efficiency scale is arbitrary.

where $L_{signal}(m, ct, \sigma_{ct})$ and $L_{background}(m, ct, \sigma_{ct})$ are the signal and background likelihood functions, respectively. Quantities N_{sig} and N_{bkg} are the yields of the signal and background events. The signal likelihood model is defined as

$$L_{signal}(m, ct, \sigma_{ct}) = \prod_{i=1}^{N} P_{Sig}(m_i) P_{Sig}(ct_i) P_{Sig}(\sigma_{ct,i}) \tag{8.12}$$

where

- $P_{Sig}(m) = \sum_{j=1}^{2} f_j \frac{1}{\sqrt{2\pi}\sigma_{m,j}} e^{-(m-\mu_m)^2/2\sigma_{m,j}}$ is a double Gaussian function that describes the B_s^0 mass peak. The standard deviations of the Gaussians are denoted as $\sigma_{m,j}$ and the average B_s^0 mass is μ_m. The relative strengths between the Gaussians are denoted as f_j and they add up to unity: $\sum_{j=1}^{2} f_j = 1$.

- $P_{Sig}(\sigma_{ct}) = \sum_{j=1}^{2} f_j \frac{\sigma_{ct}^{\gamma_j - 1} e^{-\sigma_{ct}/\beta_j}}{\Gamma(\gamma_j)\beta_j^{\gamma_j}}$ is a sum of two Gamma functions that model the decay time uncertainty distribution. The shape and scale parameters of the gamma functions are denoted as β_j and γ_j. Parameters f_j describe relative strengths between the functions.

- $P_{Sig}(ct) = \frac{1}{2c\tau_{eff}}\{e^{-ct/c\tau_{eff}} \otimes R(ct)\} \cdot \epsilon(ct)$ is an exponential decay rate function convoluted with a Gaussian ct resolution function $R(ct)$ and multiplied with a ct efficiency function $\epsilon(ct)$. The parameter $c\tau_{eff}$ is the effective lifetime of the B_s^0 meson. The resolution function is defined as $R(ct) = \frac{1}{\sqrt{2\pi}\sigma_{ct}} e^{-(ct)^2/2\sigma_{ct}}$, where the width σ_{ct} is taken from the ct uncertainty distribution in an event-by-event fashion. The efficiency function is defined as $\epsilon(ct) = k_0 + ct^{-k_1} + k_2/(ct^2 + k_3)$.

  The background component of total likelihood function is defined as

$$L_{background}(m, ct, \sigma_{ct}) = \prod_{i=1}^{N} P_{Bkg}(m_i) P_{Bkg}(ct_i) P_{Bkg}(\sigma_{ct,i}) \tag{8.13}$$

where

- $P_{Bkg}(ct) = [\sum_{k=1}^{2} f_k e^{-ct/c\tau_{bkg,k}}] \otimes R_{Bkg}(ct)$ is a sum of two exponential functions convolved with a delta function $R_{Bkg}(ct)$. Parameters $c\tau_{bkg,k}$ are the lifetimes associated to the background component of the ct distribution. Parameters f_k are the relative strengths between the exponentials and add up to unity, $\sum_{k=1}^{2} f_k = 1$.

- $P_{Bkg}(\sigma_{ct})$ is a sum of two Gamma functions.

- $P_{Bkg}(m) = e^{-Cm}$ is an exponential function that describes the background component of the mass distribution. Coefficient C is the slope of the exponential.

8.9.1 Fit procedure

A sequential unbinned extended maximum-likelihood fit to the data is performed to measure the effective lifetime of the B_s^0 meson. First, the B_s^0 mass distribution is fitted with a sum of a double Gaussian and an exponential function. The signal and the background histograms of ct uncertainty distribution are extracted from

the sidebands of the mass distribution, $5.24 < m < 5.28$ GeV and $5.45 < m < 5.49$ GeV, using the sideband subtraction method. The sum of two Gamma functions is fitted to a sideband subtracted signal component of *ct* uncertainty distribution and another sum of two Gamma functions is fitted to the background component. The values of fit parameters obtained using the individual mass fit and the fit to the components of *ct* uncertainty distributions are used as an input to two-dimensional mass and *ct* uncertainty fit. After the two-dimensional fit, the parameters of the signal and background components of the *ct* uncertainty pdfs are set constant.

The background component of *ct* distribution is extracted using the sidebands of the mass distribution. The background *ct* distribution is fitted with a sum of two exponential pdfs to obtain good input values of the background lifetimes for the last fit. Finally, a fit to the mass, *ct* and *ct* uncertainty distributions is performed to measure the effective lifetime of the B_s^0 meson.

8.9.2 Closure tests of the likelihood function

The simulated B_s^0 sample with $\Delta\Gamma_s = 0.1$ ps^{-1} is randomly divided into two independent subsamples, one containing 20% of the simulated events ($N_{events} = 85500$) and the other containing the remaining 80% of the total sample ($N_{events} = 342000$). The *ct* efficiency function is determined from the bigger subsample. The likelihood model including the new *ct* efficiency function is fitted to the smaller sample. The lifetime measured from the smaller sample is 428.3 ± 1.5 μm which is consistent with the prediction (428.4 μm). Additionally, the efficiency function is determined using the full simulated B_s^0 sample and the signal likelihood is fitted to the same sample. The fit result, 428.6 ± 0.7 μm, is also consistent with the predicted value. Based on these tests, the likelihood function is fitted to the data. The consistency test for the signal-background likelihood model is presented in Section 8.11.1 as a part of the systematic uncertainties.

	Fit result
$c\tau_{eff}$	445.2 ± 2.0 μm
$c\tau_{bkg,1}$	104.1 ± 2.0 μm
$c\tau_{bkg,2}$	475.0 ± 7.2 μm
N_{sig}	69100 ± 350
N_{bkg}	46900 ± 320
B_s^0 mean mass	5366.4 ± 0.7 MeV

Table 8.15: Best fit values of the most important parameters of the likelihood function.

8.10 Fit results

A sequential unbinned extended maximum-likelihood fit is performed to the data. The best fit values of the most important parameters of the likelihood function are listed in Table 8.15. The exact value of the measured effective lifetime is 445.24 ± 1.99 μm and it is referred as the nominal result in Section 8.11. The projections (m, ct, σ_{ct}) of the model pdf are shown in Figure 8.11. The fit projections and their pulls in the low sideband ($5.24 < m < 5.32$ GeV), high sideband ($5.32 < m < 5.41$ GeV), and signal mass regions ($5.41 < m < 5.50$ GeV) are presented in Figure 8.12. The χ^2 values and the numbers of degrees of freedom (dof) of the fit projections are shown in Table 8.16. No biases are observed in any of the fit projections or in the pull distributions.

	χ^2	dof	χ^2/dof
B_s^0 ct projection	28.4	25	1.14
B_s^0 mass projection	107.5	73	1.47

Table 8.16: χ^2 and number of degrees of freedom (dof) for the ct and mass projections of multidimensional likelihood function.

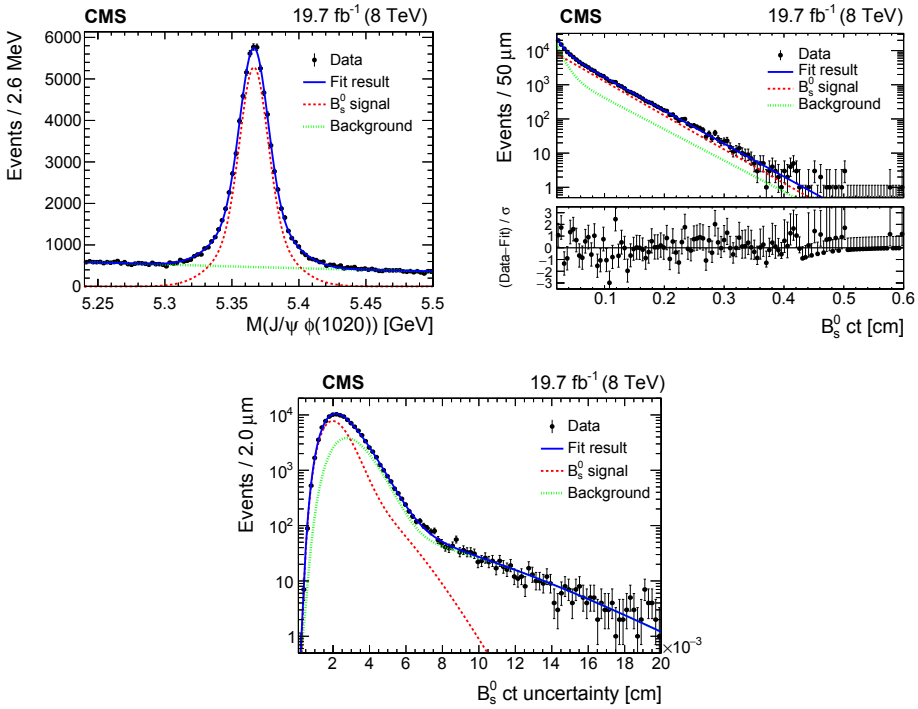


Figure 8.11: Invariant mass (left), *ct* (right), and *ct* uncertainty (bottom) distributions for B_s^0 candidates reconstructed from $J/\psi\phi$ decays. The curves are projections of the fit to the data, with the contributions from signal (dashed red line), background (dotted green line), and the sum of signal and background (solid blue line) shown.

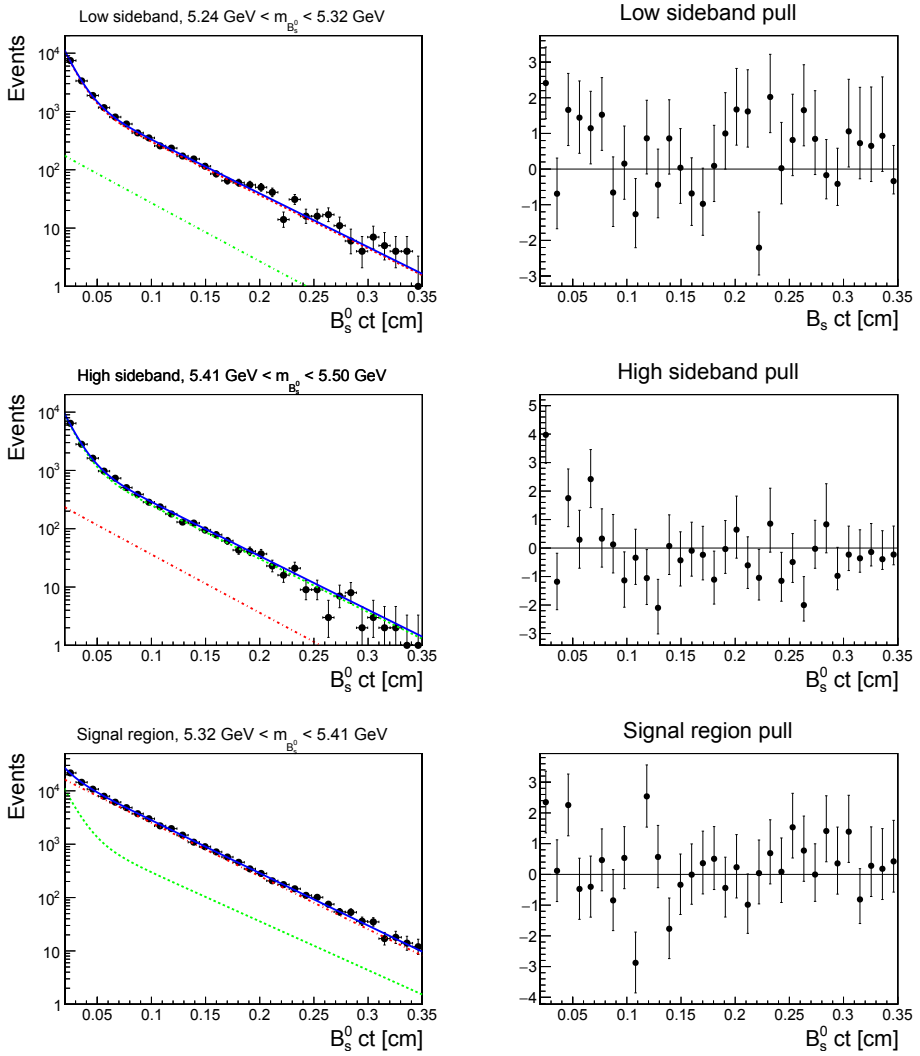


Figure 8.12: Fit projections onto *ct* axis in three B_s^0 mass regions: low sideband region 5.24 $< m_{B_s^0} <$ 5.32 GeV, signal region 5.32 $< m_{B_s^0} <$ 5.41 GeV and high sideband region 5.41 $< m_{B_s^0} <$ 5.5 GeV. The solid blue line represents the total fit, the red dashed line is the signal function, and the green dashed-dotted line is the background function.

8.11 Systematic uncertainties

Various sources of systematic uncertainties are considered to possibly affect the effective lifetime. All the contributions of systematic uncertainties are shown in Table 8.24. The largest systematic uncertainty originates from the primary vertex selection.

The fit result of the high-statistic reference mode $B^+ \to J/\psi K^+$ is also verified to be stable under different selection requirements in terms of the quality of the tracks and vertices, the kinematic variables, as well as in detector regions and data-taking periods [102]. The results are assumed to hold for the other b-hadron decay modes too, such as $B_s^0 \to J/\psi\phi$. The effect of changing the mass of the B_s^0 meson in the ct definition of Equation (8.1) from the world-average to the reconstructed mass is found to be negligible compared to other systematic uncertainties.

8.11.1 Likelihood function bias

A set of pseudo-experiments is performed to ensure that the likelihood function used in the lifetime fit does not introduce a bias in the measurement. The pseudo-data are generated and fitted with the likelihood function used in the nominal fit, and a pull distribution is formed. The pull distribution is shown in Figure 8.13. No significant bias is observed, and hence no systematic uncertainty is assigned.

8.11.2 Modelling assumptions

Biases related to the shapes of the signal and background pdfs modelling the observable distributions are evaluated by changing the signal and background pdfs separately, and then using the new likelihood function to fit the data. The residual difference between the effective lifetime results of the nominal and alternative likelihood functions is taken as the systematic uncertainty. The mass background

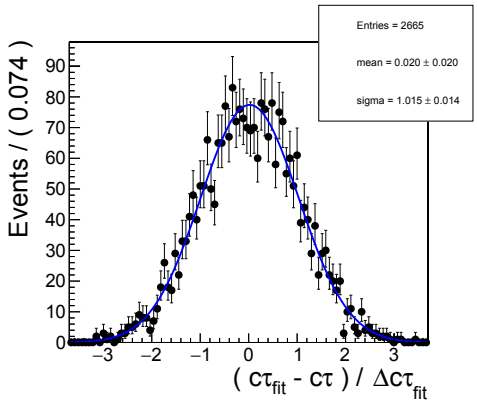


Figure 8.13: Pull distribution of the bias test of the likelihood function.

pdf is changed from an exponential function to a Chebychev polynomial. The B_s^0 mass signal pdf is switched from a double Gaussian function to a triple Gaussian function. The background component of the *ct* distribution is modelled with a triple exponential function instead of a nominal double exponential pdf. The signal and background pdfs modelling the *ct* uncertainty distribution are changed from the sums of two gamma functions to sums of a gamma function and a crystal ball function.

The fit results of the new likelihood functions ($c\tau_{eff}^{new}$) and the associated systematic uncertainty (σ_{syst}) are shown in Table 8.17. The biggest systematic uncertainty of the pdf modelling originates from the mass background model.

8.11.3 Primary vertex selection

The proton-proton collisions at the LHC contain multiple primary vertices, in which one is selected for the B hadron production vertex. The primary vertex having the smallest pointing angle between the three-dimensional flight direction and B_s^0 momentum with a constraint from a beam spot position is used as the nominal PV. The nominal PV selection is further modified, releasing the BS constraint. The

	$c\tau_{eff}^{new}$ [μm]	σ_{syst} [μm]
Mass signal model	445.23 ± 1.97	0.01
Mass background model	445.65 ± 1.98	0.41
ct background model	445.21 ± 1.99	0.03
Signal model for ct uncertainty	445.24 ± 1.98	0.0
Background model for ct uncertainty	445.24 ± 1.98	0.0
Total systematics	-	0.41

Table 8.17: Fit results ($c\tau_{eff}^{new}$) with the modified likelihood functions and the associated systematic uncertainties (σ_{syst}).

beam spot and the vertex with the largest sum of track momenta are also used as alternative vertices. The PV candidates used to quantify the systematic uncertainty are listed below.

- Beam spot (2D)
- 3D pointing angle PV without additional constraints
- PV with highest sum of track momenta
- PV with highest sum of track momenta with the beam spot constraint

Since $B^+ \to J/\psi K^+$ decay mode has the largest number of reconstructed events, the uncertainty related to PV selection is taken from B^+ lifetime fit [102]. The maximum difference between the nominal measured lifetime and the B^+ lifetimes estimated using the alternative PV choices is taken as the systematic uncertainty.

8.11.4 ct resolution

The event-by-event widths $\sigma_{ct,i}$ of the resolution function $R(ct)$ obtained from the ct uncertainty distribution are multiplied by a scale factor s, $\sigma_{ct,i} \to s \cdot \sigma_{ct,i}$, to

take into account possible over- and underestimations of the ct uncertainty value. The fit is repeated with $s = 0.5$ and $s = 2.0$. The results are shown in Table 8.18. The maximum variation with respect to the nominal lifetime result is taken as the systematic uncertainty.

	$c\tau_{eff}$ [μm]
$s = 0.5$	445.25 ± 1.97
$s = 2.0$	445.12 ± 2.06

Table 8.18: Lifetime measurements obtained by varying the scale factor s of event-by-event widths in the ct resolution function.

8.11.5 MC sample size

The number of events in the B_s^0 simulation determines the uncertainties of the efficiency parameters. In total 1000 efficiency curves are generated by varying the parameter values. The coefficients of the efficiency function are sampled using a multivariate Gaussian pdf that is constructed from the covariance matrix of the efficiency fit. The effective lifetime is fitted 1000 times using the generated efficiency functions. The distribution of the measured lifetimes is fitted with a Gaussian, whose width is taken as the systematic uncertainty associated with the finite size of the simulated sample. The efficiency curves are presented in Figure 8.14.

8.11.6 Efficiency modelling

In addition to the nominal efficiency function, the efficiency is parametrised using two alternative functions, a polynomial and a power function:

$$\begin{aligned} \epsilon_{tot}(ct) &= k_0 + k_1 ct + k_2 (ct)^2 + k_3 (ct)^3 \\ \epsilon_{tot}(ct) &= k_0 + k_1 \sqrt{ct} \end{aligned}$$

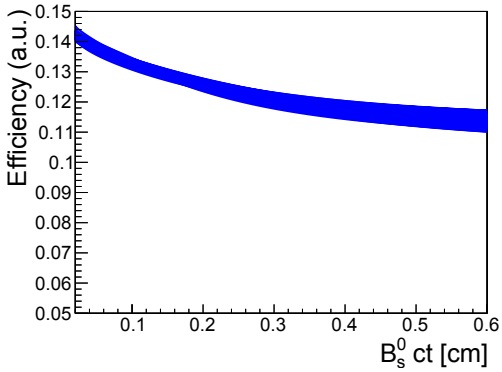


Figure 8.14: The efficiency curves used to evaluate the systematic uncertainty associated with the finite size of the simulated sample.

The alternative functions are selected such that they have similar χ^2/dof value ($\sim$ 0.9) as the nominal function. The new likelihood models are fitted to the data, and the maximum difference of the new fit result and nominal result is taken as a systematic uncertainty. The new efficiency parametrisations are shown in Figure 8.15. The fit results are shown in Table 8.19.

	$c\tau_{eff}$ [μm]
Alternative efficiency, polynomial	445.31 ± 2.06
Alternative efficiency, power function	445.08 ± 2.06

Table 8.19: The effective lifetimes obtained using the alternative efficiency functions in the likelihood model.

8.11.7 Tracker alignment

The momentum and the flight distance of the B_s^0 mesons are measured using the tracker. Therefore, misaligned tracker modules could distort the measurements of these quantities and hence introduce a systematic uncertainty into the lifetime anal-

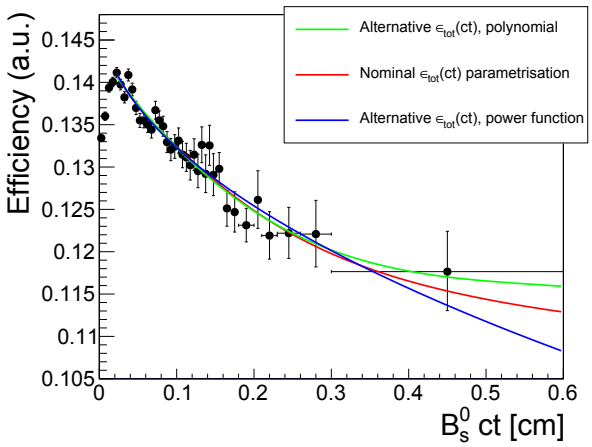


Figure 8.15: Alternative efficiency functions shown with solid blue and green lines and the nominal efficiency shown with a solid red line.

ysis. The uncertainty related to the tracker alignment is assumed to be the same within different b-hadron decays that have the same decay topology. The alignment studies are done only for the $B_d^0 \to J/\psi K^{0*}$ mode and are assumed to apply for the $B_s^0 \to J/\psi\phi$ channel too. The systematic uncertainty is evaluated with simulated events.

Nine systematic distortions [105] of the tracker module positions are assigned for the tracker geometry as a function of cylindrical coordinates ϕ, R, and z. Each misalignment scenario shifts at least one of the position coordinates of the tracker modules. For instance, the radial scaling changes the radial distance of the modules such that the new distance becomes $R' = R + c_1 R$, where the coefficient c_1 controls the shift. Also, the module positions can be squeezed or stretched as a function of the module distance from the beam axis, $z' = z + c_1 R$. Furthermore, module positions can be twisted such that $\phi' = \phi + c_1 z$.

The full B_d^0 analysis chain is rerun for all the misalignment scenarios and its lifetime is measured. The largest deviation from the nominal B_d^0 lifetime arises from the twisted scenario and is 0.5 μm. The other misalignment scenarios distort

the lifetime by 0-0.4 μm. The standard deviation of the lifetimes measured from the distorted geometries and from the nominal alignment scenario is taken as the systematic uncertainty.

8.11.8 ct range

The measured effective lifetime has to be corrected for the ct cut bias as the $ct > 200$ μm requirement distorts the fractions of heavy and light mass eigenstates in the B_s^0 sample. The correction and its uncertainty are quantified analytically. The correction is given as

$$\delta_{c\tau} = c\tau_{eff}^{cut} - c\tau_{eff} = \frac{(1-|A_\perp|^2)(c\tau_L)^2 e^{-a/c\tau_L} + |A_\perp|^2 (c\tau_H)^2 e^{-a/c\tau_H}}{(1-|A_\perp|^2) c\tau_L e^{-a/c\tau_L} + |A_\perp|^2 c\tau_H e^{-a/c\tau_H}} - \frac{(1-|A_\perp|^2)(c\tau_L)^2 + |A_\perp|^2 (c\tau_H)^2}{(1-|A_\perp|^2) c\tau_L + |A_\perp|^2 c\tau_H}, \tag{8.14}$$

where the first term represents the effective lifetime in the presence of a $ct > a$ cut defined in Equation (2.97) and the latter term is the unbiased effective lifetime defined in Equation (2.94). Additionally, the magnitude of the CP-even amplitude is given as $|A|^2 = 1 - |A_\perp|^2$ because of the normalisation constraints.

The world-average values [2] for $c\tau_H = 482.7 \pm 3.6$ μm, $c\tau_L = 426.3 \pm 2.4$ μm, and $|A_\perp|^2 = 0.250 \pm 0.006$ are used to obtain the correction $\delta_{c\tau} = 0.62 \pm 0.10$ μm, whose uncertainty is taken as the systematic uncertainty. As the ct cut enlarges the measured effective lifetime, the correction is subtracted from the fit result.

8.11.9 S-wave contribution

The B_s^0 candidates reconstructed in the $J/\psi\phi$ mode contain a small fraction of CP-odd (S-wave) $B_s \to J/\psi K^+K^-$ decay where the invariant mass of the two kaons appears inside the signal window of the $\phi(1020)$ mass.

A bias due to this so-called S-wave contribution is quantified by generating two sets of pseudo-experiments. The ct distributions of heavy and light B_s^0 states are

generated using the world-average values of the lifetimes and the amplitude $|A_\perp|^2$. In the first set of pseudo-experiments, the S-wave fraction is set to zero, but the ct cut bias is taken into account.

In the second experiment, the fraction f_H of the heavy mass state is increased by taking into account both the ct cut and the S-wave component, as in Equation (2.98). The S-wave fraction used in the pseudo-experiments is obtained by scaling the S-wave fraction measured in the ϕ_s analysis and propagating its statistical uncertainty to the generation of pseudo-datasets. Using these sets of pseudo-experiments, the correction and the systematic uncertainty are determined from the distributions of the fitted effective lifetimes.

The correction $\delta_{c\tau,S}$ to the effective lifetime is defined as the difference between the mean values of the fitted lifetimes in two distributions,

$$\delta_{c\tau,S} = \mu_{c\tau_{eff}^{cut,S}} - \mu_{c\tau_{eff}^{cut}}, \tag{8.15}$$

where $\mu_{c\tau_{eff}^{cut,S}}$ is the mean effective lifetime when the S-wave state is taken into account in the generation. Parameter $\mu_{c\tau_{eff}^{cut}}$ is the mean effective lifetime when the S-wave fraction is set to zero.

The systematic uncertainty $\sigma_{syst,S}$ is defined as the quadratic difference of standard deviations of the two effective lifetime distributions

$$\sigma_{syst,S} = \sqrt{\sigma^2_{c\tau_{eff}^{cut,S}} - \sigma^2_{c\tau_{eff}^{cut}}}, \tag{8.16}$$

where $\sigma_{c\tau_{eff}^{cut,S}}$ is the width of the distribution with non-zero S-wave fraction and $\sigma_{c\tau_{eff}^{cut}}$ is the width of the distribution where the fraction is set to zero.

The fraction of S-wave component among the signal events is found to be $f_S^{\phi_s} = N_S/(N_S + N_P) = 1.2^{+0.9}_{-0.7}\%$ in the ϕ_s analysis [12]. However, the S-wave component in the effective lifetime measurement can be different from $1.2^{+0.9}_{-0.7}\%$, since the lifetime analysis utilises different high-level trigger and signal selection compared to the ϕ_s measurement.

	N_P	N_S	$f_S = N_S/(N_S + N_P)$
ϕ_s signal selection	48707 ± 221	340 ± 18	$(6.9 \pm 0.4) \cdot 10^{-3}$
effective lifetime signal selection	75441 ± 275	648 ± 25	$(8.5 \pm 0.3) \cdot 10^{-3}$

Table 8.20: Numbers of B_s^0 candidates reconstructed in the $B_s^0 \to J/\psi\phi$ (N_P) and $B_s \to J/\psi K^+K^-$ (N_S) modes and measured from the simulated inclusive B_s^0 meson sample.

	inclusive sample	world-average
$BR(B_s^0 \to J/\psi\phi)$	0.1547	$(1.07 \pm 0.08) \cdot 10^{-3}$
$BR(B_s \to J/\psi K^+K^-)$	0.0809	$(7.9 \pm 0.7) \cdot 10^{-4}$

Table 8.21: Branching fractions of $B_s^0 \to J/\psi\phi$ and $B_s \to J/\psi K^+K^-$ decays in the inclusive B_s^0 sample and their corresponding word-average values.

The S-wave fraction present the effective lifetime analysis is evaluated using the simulated inclusive B_s^0 sample, since it contains both decay modes $B_s^0 \to J/\psi\phi$ and $B_s \to J/\psi K^+K^-$. The yields of P-wave and S-wave events (N_P, N_S) are measured with the triggers and signal selections used in the ϕ_s and effective lifetime measurements. The yields of the S- and P-wave decays are presented in Table 8.20.

Since the branching fractions of $B_s^0 \to J/\psi\phi$ and $B_s \to J/\psi K^+K^-$ modes in the inclusive sample differ from their world-average values, the numbers of S- and P-wave events reported in Table 8.20 are scaled to the numbers of events expected in the data. The branching ratios are shown in Table 8.21. The scale factors are determined by Equations (8.17)- (8.18).

$$r^P_{scale} = \frac{BR(B_s \to J/\psi\phi)_{PDG}}{BR(B_s \to J/\psi\phi)_{MC}} = (6.9 \pm 0.5) \cdot 10^{-3} \tag{8.17}$$

$$r^S_{scale} = \frac{BR(B_s \to J/\psi KK)_{PDG}}{BR(B_s \to J/\psi KK)_{MC}} = (9.8 \pm 0.9) \cdot 10^{-3} \tag{8.18}$$

The scaled S-wave fraction is

$$f_S^{scaled} = (r^S_{scale} \cdot N_S)/(r^S_{scale} \cdot N_S + r^P_{scale} \cdot N_P). \tag{8.19}$$

	f_S^{scaled}
ϕ_s signal selection	$9.8 \cdot 10^{-3}$
effective lifetime signal selection	$1.20 \cdot 10^{-2}$
ratio of S-wave fractions, $r = f_S^{scaled,c\tau_{eff}} / f_S^{scaled,\phi_s}$	1.23

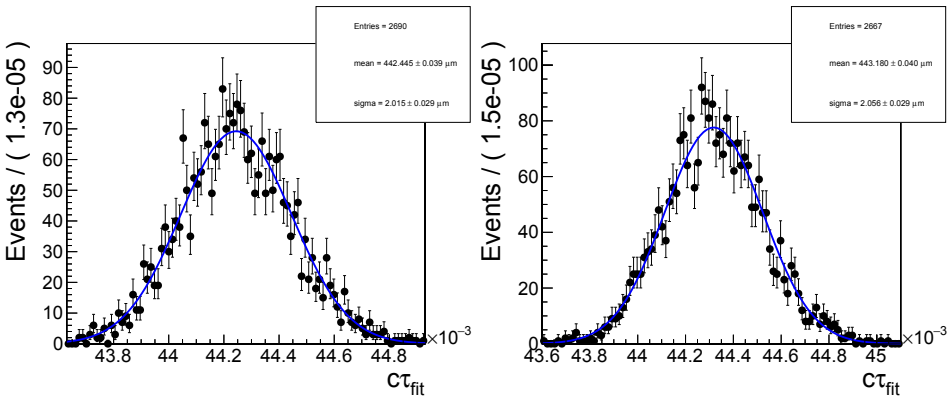
Table 8.22: Central values of the scaled S-wave fractions measured from the inclusive B_s^0 sample using the signal selection in ϕ_s and effective lifetime analyses.

The central values of the scaled S-wave fractions and their ratio are shown in Table 8.22. The ratio r is used to scale the S-wave fraction measured in the ϕ_s analysis, $f_S^{c\tau_{eff}} = r f_S^{\phi_s}$. The resulting S-wave fraction in the effective lifetime analysis is evaluated to be $f_S^{c\tau_{eff}} = 1.5^{+1.1}_{-0.9}$ %. The uncertainty of the measured S-wave fraction is of the order of 60 - 75 % of the central value and is much larger than the uncertainty associated with scale factor r. Thus, only the statistical uncertainty of the measured $f_S^{\phi_s}$ is propagated to the generation of the pseudo-experiments.

The two distributions of fitted lifetimes are shown in Figure 8.16 and the characteristics of the distributions are presented in Table 8.23. The S-wave correction is 0.74 μm, and the corresponding systematic uncertainty is 0.41 μm. Since the S-wave state increases the effective lifetime by 0.74 μm, this value is subtracted from the measured effective lifetime.

	mean effective lifetime [μm]	standard deviation σ [μm]
distribution with $f_S = 0$	442.45 ± 0.04	2.015 ± 0.029
distribution with $f_S = 1.5^{+1.1}_{-0.9}$ %	443.18 ± 0.04	2.056 ± 0.029

Table 8.23: Characteristics of the distributions presented in Figure 8.16.



(a) **(b)**

Figure 8.16: Distributions of fitted effective lifetimes. The S-wave fraction set to a) $f_S = 0\%$ and b) $f_S = 1.5^{+1.1}_{-0.9}\%$.

8.12 Results

The effective lifetime of the B^0_s meson measured from $J/\psi\phi$ decay mode using the 2012 CMS data is found to be

$$\begin{aligned} c\tau_{eff} &= 443.9 \pm 2.0 \pm 1.2\ \mu\text{m} \\ \tau_{eff} &= 1.481 \pm 0.007 \pm 0.004\ \text{ps}, \end{aligned}$$

where the first uncertainty is statistical and the second is systematic. The measured value has been corrected for the effects of ct range and S-wave contamination described in Section 8.11. The result is consistent with the measurements from other experiments and the world-average value $1.479 \pm 0.012\ \text{ps}^{-1}$ [2].

Systematic uncertainty	[μm]
MC sample size	0.6
Primary vertex selection	0.7
Tracker alignment	0.3
Likelihood function bias	-
Modelling of ct distribution	-
Modelling of ct uncertainty distribution	-
Modelling of the mass distribution	0.4
Efficiency modelling	0.2
ct resolution	0.1
ct range	0.1
S-wave contribution	0.4
Total systematic uncertainty	1.2

Table 8.24: Sources of systematic uncertainties in the B_s^0 effective lifetime measurement. If no value is reported, the systematic uncertainty is negligible compared to the statistical and other systematic uncertainties. The total systematic uncertainty is the sum in quadrature of all uncertainties.

8.13 Summary and outlook of the effective lifetime analysis

In this analysis, the effective lifetime of B_s^0 meson decaying into the $J/\psi\phi$ state was measured using 19.7 fb^{-1} of data collected at the center-of-mass energy of 8 TeV. In total, 69100 $B_s^0 \to J/\psi\phi$ signal candidates were selected from the data. The measured effective lifetime, $c\tau_{eff} = 443.9 \pm 2.0 \pm 1.2$ μm, is in agreement with the results from other experiments and with the world-average value. The precision of the effective lifetime measurement is comparable to the precision obtained from the other experiments, showing a good capability of the CMS experiment in the lifetime

analyses.

Although more sophisticated analyses exist for measuring the lifetimes of the B_s^0 eigenstates, the measurement of the B_s^0 effective lifetime gives the CMS experiment valuable information about the experimental effects affecting the lifetime measurements of B_s^0 mesons. In this analysis, we have quantified the effects of the reconstruction and selection efficiency as well as apparent efficiency on the effective lifetime, which has not been done earlier in the CMS collaboration. Furthermore, the analytical and experimental techniques developed to quantify the contamination of the additional heavy or light B_s^0 states in the lifetime measurements can be useful for the evaluation of the systematic uncertainties in future measurements. Together with the weak phase measurement, the effective lifetime analysis completes our understanding of the B_s^0 system. This provides a solid basis for the forthcoming ϕ_s and lifetime measurements for example in the $B_s^0 \to J/\psi\phi$ or $B_s^0 \to J/\psi f_0(980)$ modes.

9. Conclusions

Two precision measurements with neutral B_s^0 mesons are discussed in this thesis, a lifetime analysis and a CP violation measurement. The main result is the effective lifetime $c\tau_{eff}$ of the B_s^0 meson decaying into $J/\psi\phi(1020)$ mode,

$$c\tau_{eff} = 443.9 \pm 2.0 \pm 1.2 \ \mu\text{m}$$

where the first uncertainty is statistical and the second one systematic. The measurement is in agreement with the results from other experiments and with the world-average value. The precision of the effective lifetime measurement is comparable to the precision obtained from the other experiments. Even if more advanced methods have been developed to measure the lifetimes of the B_s^0 eigenstates, this measurement gives the CMS experiment important information about the experimental effects affecting the lifetime measurements of the B_s^0 mesons.

CP violation studies presented in the thesis concentrate on the weak phase measurement and the development and validation of the flavour tagging algorithms used to improve the precision of the ϕ_s result. The performance of the neural network-based flavour tagging algorithm is

$$\begin{aligned} \omega &= (30.2 \pm 0.2 \pm 0.05)\% \\ \epsilon_{tag} &= (8.31 \pm 0.03)\% \\ P_{tag} &= (1.307 \pm 0.031 \pm 0.007)\%, \end{aligned}$$

where the first uncertainty is statistical and the second one systematic. Using the

NN-based tagging tool, the weak mixing phase and the decay width difference were measured to be

$$\begin{aligned} \phi_{\mathrm{s}} &= -0.075 \pm 0.097 \pm 0.031 \text{ rad} \\ \Delta\Gamma_s &= 0.095 \pm 0.013 \pm 0.007 \text{ ps}^{-1}, \end{aligned}$$

where the first uncertainty is statistical and the second one systematic. The measurements of ϕ_s and $\Delta\Gamma_s$ are in agreement with the results from other experiments and the SM predictions. The CMS contribution improves the precision of the combined results of these quantities and enables further probing the standard model.

A. Data-simulation comparisons

Data-simulation comparisons of the lepton observables used in the optimisation of the cut-based tagging algorithm are shown in Figures A.1 - A.2. The signal component of the data is obtained using the Splot technique. The $B_s^0 \to J/\psi\phi$ mass distribution is used to differentiate the signal and background components. The variables presented in the figures are used in the optimisation of the cut-based tagging algorithm. The data and simulated distributions are in agreement, although a small discrepancy between the data and simulated samples is observed in the lepton ΔR distributions.

Data-simulation comparisons of the observables constructing the B_s^0 ct and ct uncertainty distributions in the effective lifetime measurement are shown in Figures A.3 - A.4. The simulated sample used in the comparisons is the one used to determine the ct efficiency, i.e., the sample with $\Delta\Gamma_s = 0.1$ ps^{-1}. The data and simulated distributions are in agreement. A small discrepancy in the L_{xy} distribution is due to a deviation between the lifetimes of the heavy and light B_s^0 states used in the simulation and measured recently.

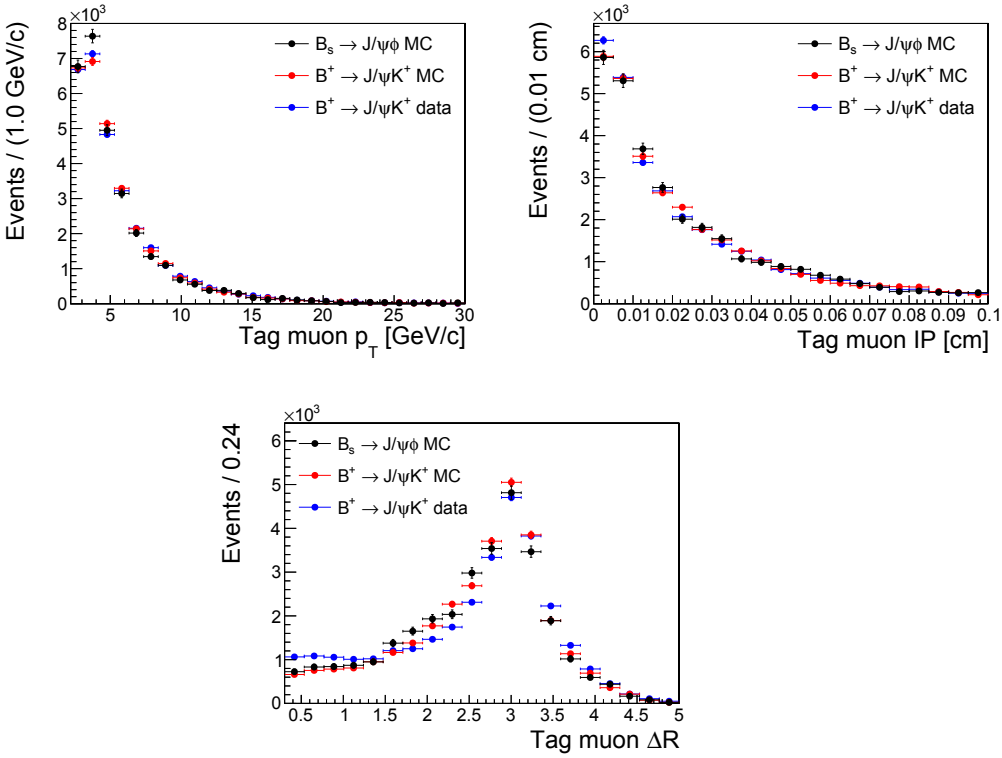


Figure A.1: Input variable comparisons for the cut-based muon tagger.

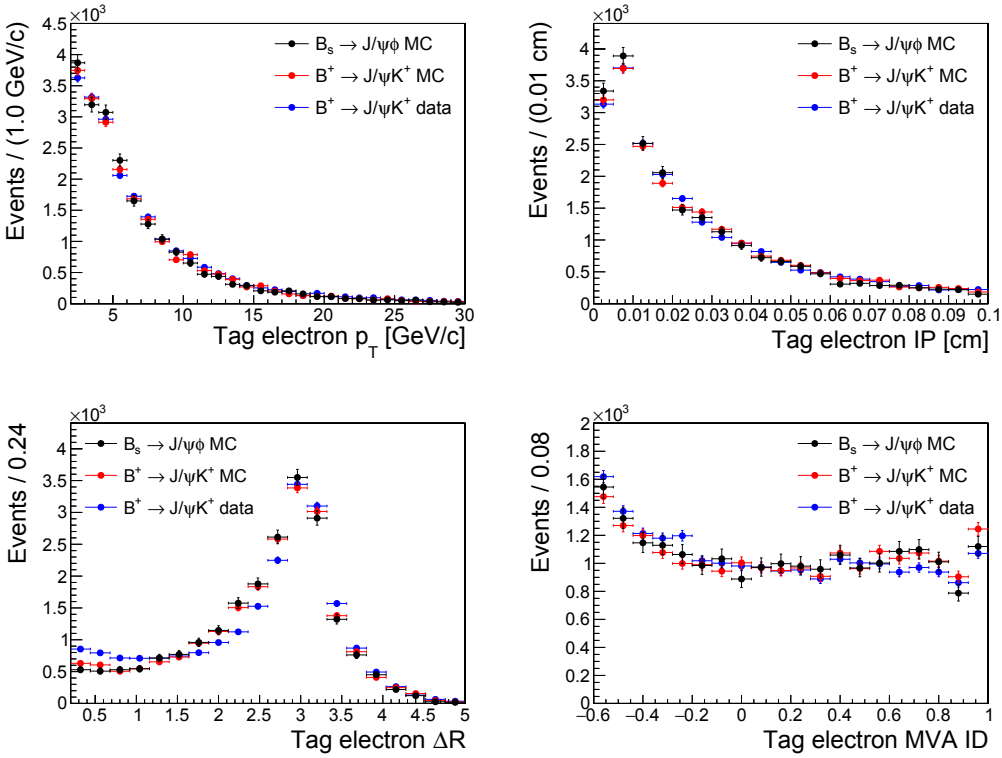


Figure A.2: Input variable comparisons for cut-based electron tagger.

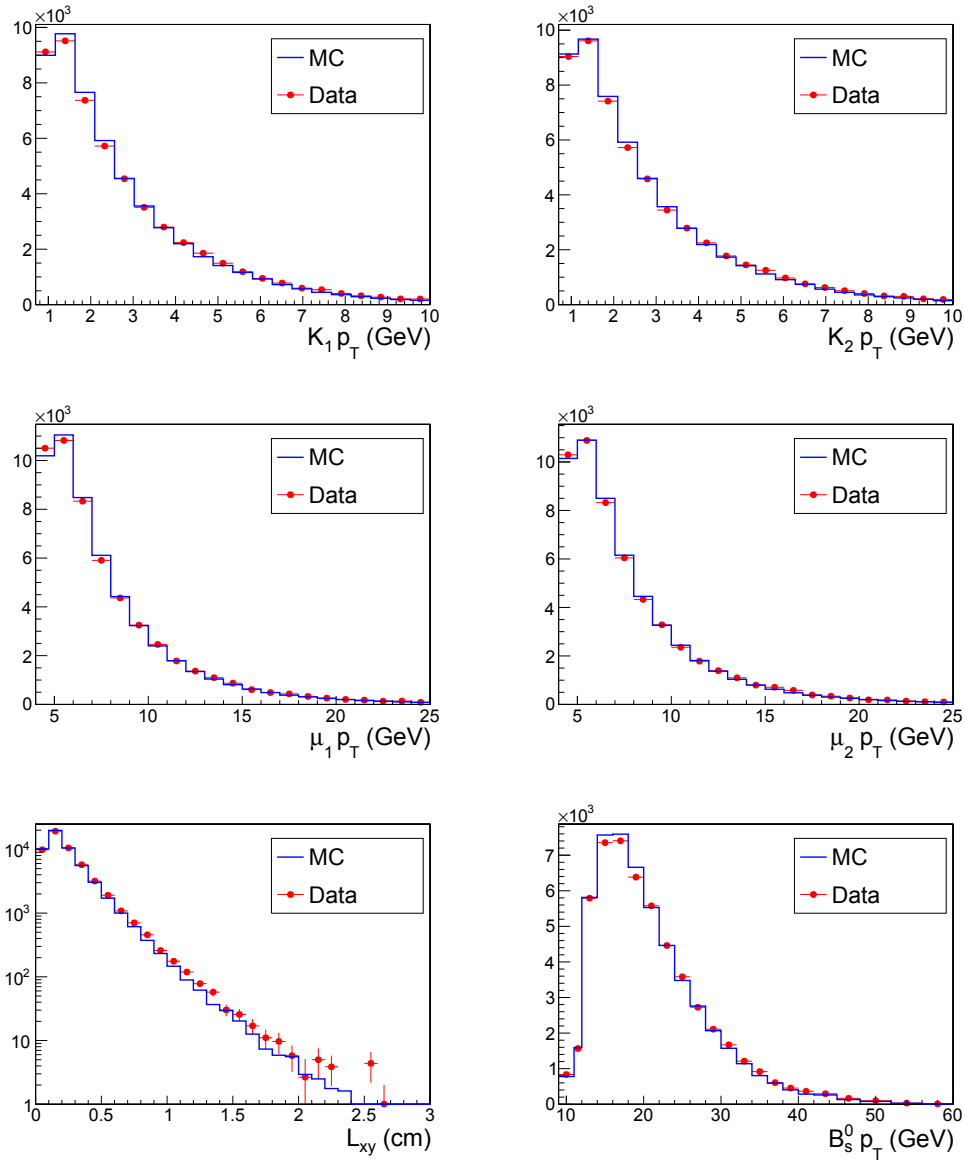


Figure A.3: Data-simulation comparisons of the key observables in the $B_s^0 \to J/\psi\phi$ effective lifetime measurement.

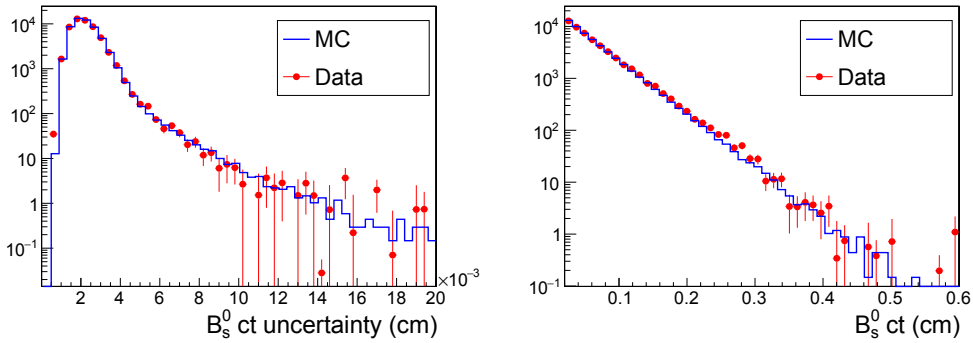


Figure A.4: Data-simulation comparisons of the key observables in the $B_s^0 \to J/\psi\phi$ effective lifetime measurement.

Bibliography

1 G. L. Bayatian et al. CMS physics: Technical design report. CERN-LHCC-2006-001, CMS-TDR-008-1, 2006.

2 C. Patrignani et al. 2016 Review of Particle Physics. *Chin. Phys. C.*, 40, 2016.

3 J.P Fernández and L. R. Flores-Castillo. Experimental insight into b-hadron lifetimes. *Progress in Particle and Nuclear Physics*, 86:1 – 34, 2016.

4 J. Charles et al. Predictions of selected flavor observables within the standard model. *Phys. Rev. D*, 84:033005, 2011. [Updated with Summer 2016 results].

5 R. Fleischer and R. Knegjens. Effective lifetimes of B_s decays and their constraints on the B_s^0-$\bar{B}_s^0$ mixing parameters. *Eur. Phys. J.*, C71:1789, 2011.

6 R. Aaij et al. Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes. *High Energy Phys J.*, 04:114, 2014.

7 F. Abe et al. Measurement of B hadron lifetimes using J/ψ final states at CDF. *Phys. Rev.*, D57:5382–5401, 1998.

8 V. M. Abazov et al. Measurement of the B_s^0 lifetime in the exclusive decay channel $B_s^0 \to J/\psi\phi$. *Phys. Rev. Lett.*, 94:042001, 2005.

9 G. Aad et al. Time-dependent angular analysis of the decay $B_s^0 \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase ϕ_s by ATLAS. *JHEP*, 12:072, 2012.

10 G. Aad et al. Flavor tagged time-dependent angular analysis of the $B_s^0 \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the weak phase ϕ_s by ATLAS. *Phys. Rev.*, D90:052007, 2014.

11 V. Khachatryan et al. Measurement of CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s \to J/\psi\phi(1020)$ decay channel. `http://cds.cern.ch/record/1744869/files/BPH-13-012-pas.pdf`, 2014.

12 V. Khachatryan et al. Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $\mathrm{B}_s^0 \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B757:97–120, 2016.

13 R. Aaij et al. Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Rev.*, D87(11):112010, 2013.

14 R. Aaij et al. Measurement of the CP-violating phase ϕ_s in $\bar{B}_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Lett.*, B736:186–195, 2014.

15 T. Aaltonen et al. First flavor-tagged determination of bounds on mixing-induced CP violation in $B_s^0 \to J/\psi\phi$ decays. *Phys. Rev. Lett.*, 100:161802, 2008.

16 T. Aaltonen et al. Measurement of the CP-violating phase $\beta_s^{J/\psi\phi}$ in $B_s^0 \to J/\psi\phi$ decays with the CDF II detector. *Phys. Rev.*, D85:072002, 2012.

17 V. M. Abazov et al. Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$. *Phys. Rev. Lett.*, 101:241801, 2008.

18 V. M. Abazov et al. Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions. *Phys. Rev.*, D85:032006, 2012.

19 M. Gell-Mann. Symmetries of baryons and mesons. *Phys. Rev. B*, 125:1067–1084, 1962.

20 H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the colour octet gluon picture. *Phys. Lett. B*, 47:365–368, 1973.

21 S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, 22:579–588, 1961.

22 A. Salam and J. C. Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168 – 171, 1964.

23 S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264 – 1266, 1967.

24 F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

25 P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.

26 P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

27 G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

28 P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.

29 T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967.

30 N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

31 M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

32 A. J. Buras and M. Lindner. *Advanced series on directions in high energy physics: Heavy flavours II.* World scientific publishing co. Pte.Ltd., 1998.

33 L. Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

34 C. Langenbruch. *Measurement of the B_s^0 mixing phase in the decay $B_s^0 \to J/\psi\phi$ with the LHCb experiment.* PhD thesis, Ruperto-Carola University of Heidelberg, 2011.

35 CKMfitter Group. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *Eur. Phys. J. C.*, 41:1–131, 2005. [Updated plots taken from: `http://ckmfitter.in2p3.fr`].

36 G. C. Branco, L. Lavoura, and J. P. Silva. *CP violation.* Oxford University Press, 1999.

37 M. Sozzi. *Discrete symmetries and CP violation: From experiment to theory.* Oxford University Press, 2008.

38 B. Adeva et al. Roadmap for selected key measurements of LHCb, 2009.

39 J. Charles et al. Current status of the standard model CKM fit and constraints on $\Delta F = 2$ new physics. *Phys. Rev.*, D91(7):073007, 2015.

40 A. S. Dighe, I. Dunietz, and R. Fleischer. Extracting CKM phases and $B_s - \bar{B}_s$ mixing parameters from angular distributions of non-leptonic B decays. *Eur. Phys. J. C*, 6:647–662. 31 p, 1998.

41 A. J. Bevan et al. The Physics of the B Factories. *Eur. Phys. J.*, C74:3026, 2014.

42 R. Aaij et al. Opposite-side flavour tagging of B mesons at the LHCb experiment. *Eur. Phys. J.*, C72:2022, 2012.

43 A.J. Buras, W. Slominski, and H. Steger. B^0 - $\bar{B}^0$ mixing, CP violation and the B-meson decay. *Nucl. Phys. B*, 245:369 – 398, 1984.

44 R. Aaij et al. Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Rev. D*, 87:112010. 39 p, 2013.

45 R. Aaij et al. Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ Decays. *Phys. Rev. Lett.*, 114:041801, Jan 2015.

46 V. M. Abazov et al. First direct two-sided bound on the B_s^0 oscillation frequency. *Phys. Rev. Lett.*, 97:021802, 2006.

47 A. Abulencia et al. Observation of $B_s^0 - \bar{B}_s^0$ Oscillations. *Phys. Rev. Lett.*, 97:242003, 2006.

48 C. Cauet et al. Optimisation and calibration of the LHCb opposite side flavour tagging. `http://indico.cern.ch/event/180683/contributions/305653/attachments/240877/337234/ccauet_PhysikE5_4526_A0.pdf`.

49 W. Herr and B Muratori. Concept of luminosity. `https://cds.cern.ch/record/941318`, 2006.

50 R. Sahoo. Relativistic kinematics. 2016. `https://inspirehep.net/record/1444984/files/arXiv:1604.02651.pdf`.

51 L. Evans and B. Bryant. LHC machine. *JINST*, 3(08):S08001, 2008.

52 J. J. Goodson and R. McCarthy. *Search for supersymmetry in states with large missing transverse momentum and three leptons including a Z-boson.* PhD thesis, Stony Brook University, May 2012.

53 O. Adriani et al. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3(08):S08006, 2008.

54 J. Pinfold et al. Technical design report of the MoEDAL experiment. Technical Report CERN-LHCC-2009-006. MoEDAL-TDR-001, Jun 2009.

55 S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3(08):S08004, 2008.

56 V. Khachatryan et al. CMS tracking performance results from early LHC operation. *Eur. Phys. J.*, C70:1165–1192, 2010.

57 S. Chatrchyan et al. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV. *JINST*, 8:P09009, 2013.

58 S. Chatrchyan et al. Performance of the CMS drift tube chambers with cosmic rays. *JINST*, 5(03):T03015, 2010.

59 F. Thyssen. Performance of the resistive plate chambers in the CMS experiment. *JINST*, 7(01):C01104, 2012.

60 S. Chatrchyan et al. Performance of the CMS cathode strip chambers with cosmic rays. *JINST*, 5:T03018, 2009.

61 S Chatrchyan et al. Performance of the CMS level-1 trigger during commissioning with cosmic ray muons. *JINST*, 5:T03002, 2010.

62 W. Adam et al. The CMS high level trigger. *Eur. Phys. J.*, C46:605–667, 2006.

63 W. Adam, B. Mangano, T. Speer, and T. Todorov. Track reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

64 R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987.

65 S. Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 3(08):P10009, 2014.

66 R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, 2007.

67 K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11):2210–2239, 1998.

68 S. Chatrchyan et al. Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

69 A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017.

70 CMS collaboration. Baseline muon selections for run-I. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId`, 2015.

71 S. Baffioni et al. Electron reconstruction in CMS. *The European Physical Journal C*, 49(4):1099–1116, 2007.

72 S. Cittolin, A. Rácz, and P. Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002.

73 K. Prokofiev and T. Speer. A kinematic and a decay chain reconstruction library, 2005.

74 J. D'Hondt et al. Fitting of event topologies with external kinematic constraints in CMS, 2006.

75 G. Fedi. *Studies of the $B_s^0 - \overline{B}_s^0$ system in the CMS experiment at the LHC*. PhD thesis, Helsinki University, 2016. [The cut-based tagging algorithm, developed by me, was first described as a part of G. Fedi's PhD thesis.].

76 J. Pazzini. *B flavour tagging with leptons and measurement of the CP violation phase ϕ_s in the $B_s^0 \to J/\psi\phi$ decay at the CMS experiment*. PhD thesis, INFN Padua, 2015. [The weighting method to determine the tagging performance of the simulated B_s and B^+ samples, developed by me and the Padova group jointly, was first described as a part of J. Pazzini's PhD thesis.].

77 T. Sjöstrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

78 David J. Lange. The EvtGen particle decay simulation package. *Nucl. Inst. Meth. A*, 462(1):152 – 155, 2001.

79 P. Golonka and Z. Was. PHOTOS monte carlo: A precision tool for QED corrections in Z and W decays. *Eur. Phys. J. C*, 45:97–107, 2006.

80 S. Agostinelli et al. Geant4 - a simulation toolkit. *Nucl. Inst. Meth. A*, 506(3):250 – 303, 2003.

81 B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Phys. Rept.*, 97:31–145, 1983.

82 G. Cerizza. *Measurements of production and decay properties of B_s mesons decaying into $J/\psi\phi$ with the CMS detector at the LHC*. PhD thesis, Tennessee University, 2012.

83 S. Chatrchyan et al. Measurement of the strange B meson production cross section with $J/\psi\phi$ decays in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D84:052008, 2011.

84 K. Nakamura et al. 2010 Review of Particle Physics. *J. Phys. G.*, 37, 2010.

85 V. Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ tev. *JINST*, 10(06):P06005, 2015.

86 V. Khachatryan et al. Measurement of $B\bar{B}$ angular correlations based on secondary vertex reconstruction at $\sqrt{s} = 7$ TeV. *JHEP*, 03:136, 2011.

87 S. Haykin. *Neural Networks : A Comprehensive Foundation*. Mcmillan College Publishing Company Inc., 1999.

88 A. Klami and M. Sjöberg. Advanced course in machine learning. `https://www.cs.helsinki.fi/courses/582744/2017/K/k/1`.

89 Jet performance in pp collisions at 7 TeV. Technical Report CMS-PAS-JME-10-003, CERN, Geneva, 2010.

90 CMS Collaboration. Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector. Technical Report CMS-PAS-PFT-10-001, CERN, 2010.

91 S. Bhattacharya. Efficiency measurement of b-tagging algorithms developed by the CMS experiment. In *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011, Providence, USA, August 9-13, 2011*, 2011.

92 M. Pivk and F. R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005.

93 CMS Collaboration. Precision lifetime measurements of b hadrons reconstructed in final states with a J/ψ meson. Technical Report CMS-PAS-BPH-13-008, CERN, Geneva, 2017.

94 R. Aaij et al. Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays. *Phys. Rev. Lett.*, 114:041801, Jan 2015.

95 G. Aad et al. Measurement of the CP-violating phase ϕ_s and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi\phi$ decays in ATLAS. *JHEP*, 08:147, 2016.

96 CMS Collaboration. Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

97 T. Järvinen. Measurement of CP violation in $B_s \to J/\psi\phi$ decays with the CMS experiment. In *16th International Conference on B-Physics at Frontier Machines (BEAUTY2016)*, 2016. https://pos.sissa.it/273/002/pdf.

98 R. Aaij et al. Determination of the sign of the decay width difference in the B_s system. *Phys. Rev. Lett.*, 108:241801, 2012.

99 J. Beringer et al. Review of particle physics. *Phys. Rev. D.*, 86:010001, 2012.

100 Y. Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016. Technical report, 2016. [Online update at http://www.slac.stanford.edu/xorg/hfag].

101 M. Artuso, G. Borissov, and A. Lenz. CP violation in the B_s^0 system. *Rev. Mod. Phys.*, 88(4):045002, 2016.

102 A. M. Sirunyan et al. Measurement of b hadron lifetimes in pp collisions at $\sqrt{s} = 8$ TeV. *Submitted to Eur. Phys. J.*, C, 2017.

103 K. A. Olive et al. 2014 Review of Particle Physics. *Chin. Phys. C.*, 38, 2014.

104 T. du Pree. *Search for a strange phase in beautiful oscillations.* PhD thesis, Vrije University, Amsterdam, 2010.

105 S. Chatrchyan et al. Alignment of the CMS tracker with LHC and cosmic ray data. *JINST*, 9:P06009, 2014.